



Les constructions syntaxiques de l'écrit scientifique : exploration et analyses de corpus

Fifielska Emilia

► To cite this version:

Fifielska Emilia. Les constructions syntaxiques de l'écrit scientifique : exploration et analyses de corpus. Linguistique. 2015. dumas-01213405

HAL Id: dumas-01213405

<https://dumas.ccsd.cnrs.fr/dumas-01213405>

Submitted on 8 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Les constructions syntaxiques de l'écrit scientifique – exploration et analyses de corpus

**FIFIELSKA
Emilia**

Sous la direction de Mme Agnès TUTIN

Laboratoire Lidilem (Didactique des Langues et Ingénierie Pédagogique
Multimédia)

UFR LLASIC
Département Sciences du langage et français langue étrangère
Mémoire de master 2 recherche - 30 crédits

Spécialité ou Parcours : Linguistique

Année universitaire 2014-2015



Les constructions syntaxiques de l'écrit scientifique – explorations et analyses de corpus

**FIFIELSKA
Emilia**

Sous la direction de Mme Agnès Tutin

Laboratoire : Lidilem (Didactique des Langues et Ingénierie Pédagogique
Multimédia)

UFR LLASIC
Département Département Sciences du langage et français langue étrangère

Mémoire de master 2 recherche - 30 crédits

Spécialité ou Parcours : linguistique

Année universitaire 2014-2015

Remerciements

Pour commencer, je tiens à exprimer ma reconnaissance à ma Directrice de mémoire, Madame Agnès Tutin pour sa grande disponibilité tout au long de la rédaction de ce mémoire. Je voudrais la remercier de m'avoir orientée vers ce sujet, pour les conseils précieux et pour la correction de mes erreurs.

Je remercie aussi Monsieur Olivier Kraif pour son aide dans les aspects informatiques de mon mémoire.

DÉCLARATION

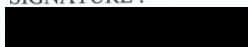
1. Ce travail est le fruit d'un travail personnel et constitue un document original.
2. Je sais que prétendre être l'auteur d'un travail écrit par une autre personne est une pratique sévèrement sanctionnée par la loi.
3. Personne d'autre que moi n'a le droit de faire valoir ce travail, en totalité ou en partie, comme le sien.
4. Les propos repris mot à mot à d'autres auteurs figurent entre guillemets (citations).
5. Les écrits sur lesquels je m'appuie dans ce mémoire sont systématiquement référencés selon un système de renvoi bibliographique clair et précis.

NOM : F. FIELSKA

PRENOM : Emilia

DATE : 04.09.2015

SIGNATURE :



Sommaire

Introduction	8
Partie 1 - L'écrit scientifique sous différentes perspectives	9
CHAPITRE 1. GENRE ET ECRITS SCIENTIFIQUES.....	120
1. VARIATION ET GENRE	120
2. LANGUE GENERALE ET LANGUE DE SPECIALITE	142
3. L'ARTICLE SCIENTIFIQUE COMME UN GENRE.....	164
CHAPITRE 2. CARACTERISTIQUES LINGUISTIQUES DE L'ECRIT SCIENTIFIQUE	19
1. GENRE ET L'ANALYSE SYNTAXIQUE.....	19
2. CARACTERISTIQUE DU DISCOURS SCIENTIFIQUE SELON DES TRAVAUX PRECEDENTS.....	21
3. CONSTRUCTIONS SYNTAXIQUES.....	276
4. LE MODELE DE CONSTRUCTION ADOPTE.....	293
Partie 2 - Méthodologie	376
CHAPITRE 3. LA RECHERCHE DES CONSTRUCTIONS	37
1. LE SCIENTEXT	37
2. LE CHOIX DES CORPUS	38
3. LE TRAITEMENT DU CORPUS	39
4. LES RESULTATS.....	40
5. LE CORPUS DE CONTRASTE.....	41
Partie 3 - Les résultats et l'observation des constructions	44
CHAPITRE 4. LES RESULTATS QUANTITATIVES.....	465
CHAPITRE 5. LA DESCRIPTION DES CONSTRUCTIONS.....	510
1. L'IMPERSONNEL.....	50
2. LE PASSIF	56
Conclusion.....	62

Introduction

Le présent mémoire s'inscrit dans les travaux du laboratoire LIDILEM de l'Université Stendhal 3 de Grenoble sur l'écrit scientifique et, plus précisément, dans la continuité des travaux du projet Scientext. Ce projet avait pour objectif principal d'étudier les marques de positionnement dans l'écrit scientifique. À cette fin, on a constitué un corpus des écrits scientifiques de français et d'anglais qui comprend plusieurs disciplines (appartenant aux sciences humaines, sciences expérimentales et sciences de l'ingénieur) et plusieurs genres (articles scientifiques, thèses, actes de colloques, écrits d'étudiants). Le corpus ainsi constitué permet aux linguistes et aux didacticiens de mener des recherches sur les caractéristiques linguistiques de l'écrit scientifique .

Selon Rinck (2010), il existe deux types d'approches dans l'étude du discours scientifique :

Les premières interrogent l'activité scientifique et montrent l'intérêt de le faire à travers les signes et les textes qu'elle produit. Les secondes visent à décrire les caractéristiques linguistiques du discours scientifique et éclairent ainsi l'activité scientifique et la manière dont elle se construit à travers ses discours. (Rinck, 2010, p. 429).

Dans notre travail, nous nous situons dans la deuxième approche : en discernant des caractéristiques de l'écrit scientifique, nous voudrions mener une réflexion sur la façon dont la langue scientifique se construit et sur les fonctions pragmatiques de l'écrit scientifique.

L'écrit scientifique est un écrit de spécialité. Or, les langues de spécialité sont d'habitude envisagées de point de vue terminologique ou lexicologique. Dans notre perspective, il s'agit de se concentrer sur l'aspect syntaxique, tout en prenant en compte les propriétés sémantiques qui y sont associées. Plus précisément, nous essayerons de repérer les constructions surreprésentées dans l'écrit scientifique grâce à l'exploitation du corpus Scientext. Ensuite, nous décrirons ces constructions en analysant leurs constituants et leur environnement distributionnel. Dans un dernier temps, nous chercherons à comprendre pourquoi ces structures apparaissent fréquemment dans l'article scientifique, en identifiant leurs fonctions pragmatiques, discursives ou stylistiques.

Ce mémoire a pour l'objectif d'observer les constructions syntaxiques particulièrement représentées dans l'écrit scientifique en vue de mieux discerner la nature

et le fonctionnement de ce type de discours. Il s'agit d'étudier les constructions syntaxiques en sciences humaines et sociales (économie, sociologie, sciences politiques, sciences de l'éducation, sciences de l'information, psychologie, linguistique, histoire, géographie, anthropologie) en nous basant sur le corpus Scientext. Notre étude peut être donc classifiée comme relevant de la syntaxe avec des analyses sémantiques et de la linguistique du corpus.

Notre motivation pour les constructions syntaxiques est liée à la nouveauté de cette approche dans ce type de discours. Les études antérieures se concentrent plutôt sur le vocabulaire et les phraséologismes de l'écrit scientifique. Ainsi, nous pouvons citer quelques travaux concernant l'aspect terminologique de l'écrit scientifique, comme Louveau et Tolas (2002) qui étudient la phraséologie de la médecine, comme Dury (2006, 2008) qui s'occupe de l'évolution diachronique du langage écologique ou des travaux aussi axés sur l'aspect lexicologique de l'écrit scientifique, mais concernant le lexique transdisciplinaire non spécialisé, fréquent dans ce genre (Pecman 2004, Tutin 2007). Cependant, nous présumons que les différences entre la langue de spécialité et la langue générale concernent tous les niveaux de la langue. Nous soutenons donc qu'il faudrait combler ce vide et initier des études syntaxiques sur le discours scientifique. Si certains travaux comme ceux de Kocourek ont déjà abordé cette question, il serait pertinent d'effectuer des études plus détaillées, et de type quantitatif, en utilisant les techniques de la linguistique de corpus. Cela peut se révéler particulièrement utile en vue des objectifs didactiques, pour l'enseignement aux étudiants locuteurs non natifs dans le cadre de la didactique du français de spécialité ou du français sur objectifs spécifiques (FOS).

Notre mémoire s'organise en cinq chapitres. Les deux premiers concernent les apports théoriques : nous réfléchirons sur des notions synonymiques telles que la variation, le registre, le genre et leur rapport à l'écrit scientifique, sur la relation entre langue non spécialisée et langue de spécialité, sur l'article scientifique vu comme un des genres, sur le rapport entre le genre et la syntaxe. Après avoir répondu à ces questions, nous relèverons quelques traits morphosyntaxiques de la langue scientifique en nous basant principalement sur les travaux de Kocourek (1982) et de Biber (2006), deux auteurs qui ont eu une réflexion sur la syntaxe dans la langue scientifique. Nous nous intéressons aussi aux approches grammaticales nécessaires pour modéliser nos données, en particulier les modèles du lexique grammaire et de la grammaire de constructions. Ils nous faciliteront également la tâche de construction de notre propre modèle d'analyse des constructions. Le

chapitre suivant concernera la méthodologie : nous décrirons brièvement le corpus choisi et les heuristiques nous permettant de mettre en évidence les propriétés des constructions. Ensuite, nous procéderons à l'analyse des constructions. La première partie sera consacrée à l'aspect quantitatif (chapitre 4) et nous indiquerons le nombre d'occurrences et le résultat relatif à la taille du corpus de chaque construction analysée. Par la suite, grâce à notre modèle, nous décrirons les constructions trouvées et nous essayerons de décrire leurs fonctions pragmatiques, communicationnelles, discursives et stylistiques (chapitre 5).

Partie 1

-

L'écrit scientifique sous différentes perspectives

Chapitre 1. Genre et écrits scientifiques

Pour notre travail sur les constructions de l'écrit scientifique, il nous faut aborder la question du discours scientifique en nous servant des termes de *variation*, *genre*, *langue de spécialité*. Chacune de ces notions fait ressortir des caractéristiques différentes de la langue scientifique, ce qui nous permettra d'aborder le sujet de notre étude sous des angles divers. Dans un premier temps, nous allons discerner les notions de *variation* et de *genre*.

1. *Variation et genre*

Pratiquement toute description linguistique comporte une variation de la langue générale, la variation se produisant lorsqu'une langue naturelle s'actualise en discours (Desmet, 2006). La notion de variation entraîne d'autres termes souvent synonymiques et ambigus comme *niveau de langage*, *genre*, *registre*, *style*. En vue de notre objectif de recherche, à savoir l'écrit scientifique, nous allons nous servir principalement du terme « genre ». Néanmoins, les autres termes susmentionnés peuvent aussi servir à caractériser la nature du discours scientifique.

La variation qui concerne les ordres stylistique ou situationnel est appelée la variation diaphasique. C'est la variation qui, selon nous, inclut les écrits scientifiques : la situation concerne le lieu de la production, les établissements de recherche dans la communication formelle ; en ce qui concerne le style, elle exige un registre soigné et de la rigueur intellectuelle qui, à leur tour, entraînent la cohérence textuelle. La variation diaphasique concerne principalement l'adaptation à la situation des locuteurs. Grâce à sa compétence, le locuteur, avec plus ou moins d'habileté et de réussite selon le cas, choisit un niveau d'expression adéquat pour utiliser ce qu'on pourrait désigner par le terme de « répertoire discursif », c'est-à-dire : « la possibilité, variable suivant les locuteurs et suivant les moments de sa vie langagière, d'utiliser les formes génériques de la communication verbale et d'en jouer en fonction de ses besoins » (Beacco, 2004, p. 110). Les phénomènes relevant de cette variation concernent plusieurs niveaux de la langue : la prononciation, le lexique, la morphologie, la syntaxe, ainsi que sa structure textuelle et les aspects pragmatiques.

La notion de variation diaphasique est souvent confondue avec celle de variation diastratique, la première étant caractérisée selon l'ordre « stylistique et situationnel » (Gadet, 1996), la deuxième selon « la dimension sociale ou démographique » (Gadet,

ibidem.), toutes les deux traitent très souvent les mêmes phénomènes linguistiques. Les liens entre le diaphasique et le diastratique peuvent être retrouvés dans le cas de la langue scientifique : l'écrit scientifique, qui entre dans la variation diaphasique, n'est pas toujours accessible aux classes sociales défavorisées (le domaine du diastratique). Il y a évidemment une relation étroite entre l'appartenance à un milieu social et la capacité à maîtriser la variation diaphasique : « plus le niveau socioculturel est placé haut dans la hiérarchie des niveaux, plus l'éventail des registres tend à s'ouvrir et à se nuancer » (Bourquin, 1965 cité par Gadet, 1996, p. 22).

La langue scientifique a un style spécifique dominé par les fonctions métalinguistiques et référentielles. Il sert principalement à la description des faits, à la démonstration de la vérité. Les recommandations concernant le style de l'écrit scientifique sont la simplicité, la clarté, la transparence. Le langage scientifique se veut économique (sans redondance ni superflu), l'information doit être pertinente et précise. C'est une logique argumentative qui structure le texte où chaque proposition construit progressivement le sens.

Pour Bakhtine (1986), la notion de style est liée à la fois au genre de discours et à l'empreinte individuelle d'un locuteur (Freyermuth S., 2012, p. 175). Il y a des genres qui sont plus enclins à traduire le caractère individuel d'un locuteur, comme les genres littéraires. De plus, Bakhtine met les faits grammaticaux et les faits stylistiques au même niveau, les premiers concernant les faits abstraits vus du point de vue de la langue, les deuxièmes, les faits concrets vus du point de vue de l'« énoncé concret », les premiers peuvent être donc seulement perceptibles si actualisés dans un énoncé singulier. Ces termes renvoient aux notions de « langue » et « parole » comprises dans le sens saussurien et, comme l'étude de la langue est indissociable de son utilisation effective, Bakhtine souligne le fait que la dimension stylistique cachée sous un terme abstrait de « genre » est présente dans tout énoncé concret : « La langue pénètre dans la vie à travers des énoncés concrets (qui la réalisent), et c'est encore à travers des énoncés concrets que la vie pénètre dans la langue. » (Bakhtine, 1986 : 268 cité d'après Freyermuth, 2012 p. 174). Ainsi, tout énoncé est façonné selon un genre de discours particulier et, à l'inverse, ce sont des énoncés qui influent sur un genre donné.

La notion de genre peut être envisagée de deux points de vue - linguistique et social. De fait, selon Bakhtine, il existe une relation d'interdépendance entre genres et des activités humaines différentes : l'utilisation de langue advient au sein des domaines

d'activités humaines ; pour cela, la langue ne peut pas être conçue séparément de domaine d'activité :

Each separate utterance is individual, of course, but each sphere in which language is used developed its own relatively stable types of the utterances. These we may call speech genres. (Bakhtine, 2004, p. 60)

Vu la multitude des champs d'activité humaine, le nombre de genres semble presque infini.

Nous voudrions aussi mentionner le caractère variable du genre qui est un « dispositif de communication, un ensemble de normes, variables dans le temps et dans l'espace, qui définissent certaines attentes de réception. » (Maingenu, 1993, p.15). Outre le changement diachronique, cette variabilité peut être comprise comme l'écart de certains traits par rapport aux traits prototypiques.

Pour notre travail, nous emploierons le terme de genre.

2. Langue générale et langue de spécialité

La définition et le statut de langue de spécialité rencontrent des traitements diversifiés, voire opposés selon les chercheurs. Dans cette section, nous allons présenter quelques-uns de ces points de vue et situerons le nôtre.

Un problème concernant la notion de la langue de spécialité est l'identification des caractéristiques qui permettraient de la séparer de la langue générale. Traditionnellement, c'était un domaine restreint dans lequel opéraient les locuteurs se servant d'une langue de spécialité qui permettait de l'identifier. Or, les sciences étant de plus en plus interdisciplinaires, un domaine spécifique ne semble plus être un indice pertinent qui identifierait une langue de spécialité. Pour notre recherche, nous nous appuyons sur la définition de Lerat (1955) selon qui « les langues de spécialité sont des langues naturelles dans leur fonction de communication de savoirs spécialisés ». Cette fonction de communication de savoirs spécialisés dirige la pensée vers une caractéristique principale de la langue de spécialité, c'est-à-dire le fait que la divergence entre l'intention de l'émetteur et la compréhension du récepteur est réduite au minimum.

En ce qui concerne le statut des langues de spécialité, nous adoptons le point de vue de Kocourek définissant la langue de spécialité comme « une sous-langue de la langue dite naturelle » (Kocourek, 1982, p.13.) Selon lui, la langue de spécialité est un des systèmes sémiotiques au sein du discours de spécialité (qui comprend aussi les ensembles signifiants non linéaires tels que figures ou graphiques et les langages artificiels, p.ex. le langage symbolique).

Contrairement à Kocourek ou Spillner (1992) qui, lui aussi, attribue à la notion de langue de spécialité des différences à tous les niveaux de la langue, certains chercheurs n'accordent pas à la langue de spécialité la place d'une sous-langue et préfèrent parler des vocabulaires ayant des « emplois particuliers du français et de ses variétés, qui font appel, pour la prononciation, la morphologie et la syntaxe, au fonds de la langue commune ». (B. Quemada, 1978, cité par Calberg-Challot, 2008, p.78).

Selon Gaudin (1995), la notion de langue de spécialité est difficile à définir sans avoir recours à d'autres termes également flous. Pour lui, la langue de spécialité est une notion de sens commun et, comme telle, elle n'est utile que dans la pratique quotidienne. Faute, selon lui, d'études comparatives qui permettraient d'identifier des différences strictement linguistiques dans la diversité du discours de spécialité, il s'oppose à cette notion. D'après lui, malgré l'existence des vocabulaires particuliers qui ont principalement la fonction identitaire d'une collectivité, la division en langues de spécialité n'est qu'un artefact qui assouvit une « manie classificatoire » (Gaudin, 1995, p.231). C'est une attitude opposée aux propositions de Lerat selon qui « la dénomination de langue de spécialité souffre d'induire une fragmentation et une marginalisation qui sont contre-intuitives » (Lerat, 1995, p. 19). C'est pour souligner une grande diversité au sein du concept *langue de spécialité* que Lerat a préféré le pluriel et le terme de *langues spécialisées*.

De notre côté, nous soutenons que la fonction communicative de la langue de spécialité, qui veut rendre chaque message le plus univoque possible, implique des choix spécifiques de tous les niveaux de la langue. Outre la fonction communicative, chaque langue de spécialité a ses propres fonctions et se concentre sur une réalité extralinguistique différente, des choix linguistiques qui se différencient d'une langue de spécialité à l'autre. C'est pourquoi nous partageons le point de vue de Lerat en ce qui concerne la pluralité de langues spécialisées et la position de Kocourek qui attribue au langage scientifique un statut d'une sous-langue en se détachant des approches limitant les langues de spécialité à des vocabulaires spécifiques.

3. L'article scientifique comme genre

Le discours scientifique appartient à un genre complexe selon la division bakhtinienne entre genres simples et genres complexes. (Freyremuth, 2011, p. 174). Les premiers se sont formés lors de l'activité orale spontanée. Ils pénètrent et modifient des genres complexes qui, à leur tour, surviennent principalement dans des activités plutôt évoluées comme les domaines artistiques, scientifiques, sociopolitiques. Ainsi, entrent dans des genres complexes, outre le discours scientifique, le conte, le roman, le théâtre, le discours idéologique, le discours du droit, le discours de la médecine, le discours journalistique.

Le schéma de Desmet (2006) divise le discours scientifique en :

- discours scientifique spécialisé (p.ex. une revue scientifique),
- discours scientifique pédagogique ou didactique (p.ex. un manuel pour les étudiants),
- discours de semi-vulgarisation scientifique (une revue de spécialité pour un public d'initiés),
- discours de vulgarisation scientifique (p.ex. les secteurs spécialisés dans la presse générale).

Le discours scientifique peut se diviser en plusieurs sous-genres qui peuvent être divisés en deux groupes principaux: les genres écrits (articles de recherches, monographies, résumés, rapports) et les genres oraux (communications de conférences). Les genres ainsi vus sont définis sur la base de l'événement communicationnel auquel ils participent. (Pic et Furmaniak, 2010).

Dans notre recherche, nous sommes intéressée au discours scientifique spécialisé, et, plus précisément, à l'article scientifique.

L'article scientifique peut être perçu comme un des moyens de communication d'une communauté de discours, à savoir une communauté qui définit les conventions selon lesquelles le discours opère ; l'utilisation de la langue dans un groupe est vue comme un comportement social et le discours au sein du groupe est un moyen de maintenir et d'étendre les connaissances de ce groupe. Le discours est aussi une partie constitutive et un identificateur du groupe. Selon Swales (1990), une communauté de discours possède six caractéristiques principales:

- a) Une communauté de discours a des objectifs publics communs qui sont partagés par ses membres. Ces objectifs peuvent être explicites, rédigés sur papier ou ils peuvent être plus tacites. C'est le partage de l'objectif et pas du sujet qui est un critère essentiel de la distinction d'une communauté de discours, même si le premier implique très souvent le second.
- b) Une communauté de discours a un système d'intercommunication entre ses membres. Ce système peut prendre des formes diverses : des rencontres, des télécommunications, de la correspondance, des newsletters, des conversations. Swales souligne l'importance de ce critère : même si les personnes opèrent dans le même milieu et concentrent leurs énergies sur les mêmes activités, les mêmes buts, on ne peut pas considérer qu'il s'agit d'une communauté de discours si elles n'interagissent entre elles.
- c) Le but principal de la participation dans une communauté de discours est d'échanger des informations. Cela permet d'assurer les objectifs secondaires d'une communauté de discours comme améliorer la performance d'une équipe de football, faire de l'argent dans une maison de courtage ou développer les moyens permettant de mieux cultiver les roses dans un club de jardinage ou, dans le cas du langage scientifique, produire des connaissances ou comprendre des phénomènes en vue des applications fonctionnelles.
- d) Une communauté de discours utilise et donc possède un ou plusieurs genres dans la poursuite de ses objectifs de communication. Un genre est créé par des attentes discursives d'une communauté de discours : des sujets pertinents, la forme, la fonction et le positionnement des éléments discursifs, le rôle qu'un texte joue au sein d'une communauté de discours.
- e) Une communauté a acquis un certain lexique spécifique. Ce lexique peut être partagé par d'autres communautés, il peut être également hautement spécialisé, mais le plus fréquemment ce phénomène passe par le développement d'abréviations et d'acronymes spécifiques. Le lexique spécifique d'une communauté assure la communication efficace entre experts.
- f) Une communauté de discours a un seuil de membres avec un degré approprié d'expertise qui fournissent un contenu pertinent et possèdent une compétence discursive. La survie d'une communauté dépend de la proportion adéquate entre les novices qui entrent comme apprentis et les experts.

L'article scientifique reste le genre du discours scientifique le plus étudié. De fait, il est le moyen d'échange le plus courant entre les chercheurs et pour cela, il est facilement accessible et abordable pour les études sur le discours scientifique écrit. C'est d'ailleurs le genre du discours scientifique proposé pour la recherche sur le discours scientifique par Swales (1990) en tant que texte hiérarchiquement organisé divisé en sections séparées. Le but principal de l'article scientifique est celui de transmettre les résultats de recherche. Le texte devrait donc être neutre et objectif. Or, on peut attribuer au discours scientifique la fonction rhétorique de convaincre et de persuader, et donc une certaine partie de subjectivité (Grossmann et Tutin, 2013, Hyland, 2009, Fløttum et al., 2006,) : l'auteur de l'écrit scientifique veut diffuser ses propos au sein des autres membres de la communauté scientifique. Cet aspect persuasif implique une autre caractéristique de l'article scientifique, à savoir sa dimension interactionnelle : l'auteur s'appuie sur des recherches antérieures et il polémique contre elles en dialoguant avec d'autres spécialistes de la branche. Outre cela, nous voyons l'aspect interactionnel dans le dialogue implicite avec le lecteur : l'auteur essaie de réfuter des reproches présumés avant qu'ils lui soient présentés; c'est ce que nous pourrions appeler le dialogisme interlocutif (Grossmann et Tutin, 2010).

Le destinataire de l'article scientifique est, le plus souvent, un autre chercheur, spécialiste dans le même domaine, du même camp. Or, l'objectif communicationnel et le destinataire agissent sur la forme et les caractéristiques de l'écrit. Pour cela, le langage utilisé se veut précis, clair, ce qui à son tour entraîne l'emploi d'une terminologie hautement spécialisée. L'objectif de base de l'article scientifique est de communiquer des idées et de diffuser les résultats d'une recherche effectuée. L'article scientifique est, donc, soumis à deux types de règles : il doit obéir aux exigences du processus de recherche et aux exigences de l'expression écrite en ce qui concerne le niveau stylistique et la structuration du texte.

Notre travail concerne le discours scientifique spécialisé écrit, à savoir l'article scientifique. Nous examinerons dans ce qui suit les caractéristiques linguistiques de l'écrit scientifique, ainsi que quelques modèles descriptifs qui nous permettront d'élaborer notre propre modèle d'analyse des constructions.

Chapitre 2. Caractéristiques linguistiques de l'écrit scientifique et modèles descriptifs.

Ce chapitre va être consacré aux questions liées à la grammaire. D'abord, nous allons aborder les liens entre le genre linguistique et l'analyse syntaxique. Ensuite, nous passerons aux travaux précédents sur la syntaxe dans le discours scientifique pour, nous concentrer ensuite sur les approches grammaticales qui nous permettront d'effectuer notre propre analyse des constructions. Nous présenterons enfin notre modèle de description des constructions.

1. Genre et l'analyse syntaxique

Dans ce chapitre, nous allons aborder le sujet de l'analyse syntaxique en fonction du genre. Nous sommes persuadée que les liens entre la grammaire et la stylistique existent, tout comme l'a déjà souligné Bakhtine :

La grammaire et la stylistique se rejoignent et se séparent dans tout fait de langue concret qui, envisagé du point de vue de la langue, est un fait de grammaire, envisagé du point de vue de l'énoncé individuel est un fait de stylistique. Rien que la sélection qu'opère le locuteur d'une forme grammaticale déterminée est déjà un acte stylistique. Ces deux points de vue sur un seul et même phénomène concret de langue ne doivent cependant pas s'exclure l'un l'autre, ils doivent se combiner organiquement (avec le maintien méthodologique de leur différence) sur la base de l'unité réelle que représente le fait de langue [...]. (Bakhtine, 1986 : 272, cité par Adam, 2011, p. 116).

Bien qu'il y ait des linguistes qui, comme Bakhtine, aient déjà attiré l'attention sur l'interdépendance des choix grammaticaux et le type de discours, les études portant sur les relations entre genres du discours et grammaire ne sont pas nombreuses. (Krazem, 2011) Même si certains syntacticiens se sont penchés sur un genre pour confirmer ou restreindre une règle grammaticale, la plupart d'entre eux ont ignoré les genres du discours. Pourtant, certains linguistes comme Krazem (2011) ont une approche différente ; dans leurs explications de certains phénomènes linguistiques, ils s'appuient très souvent sur les genres, au point d'en utiliser les termes, comme dans : présent de narration, présent sportif, infinitif de prescription, imparfait forain (Krazem, 2011). Les linguistes de la tradition bakhtinienne ont estimé que ce sont les faits de langues, à savoir les points de grammaire, qui sont indispensables à la description du genre. Selon eux, l'analyse des genres dépourvue de son aspect grammatical relève plutôt de la science de la communication, de la sociologie que de la linguistique.

Les langues de spécialité ne constituent pas un système grammatical à part, mais ils utilisent celui de langue générale. Entre les deux, ce n'est pas une rupture, mais un continuum (Lerat, 1955 : 20). Comme l'exposent Sager, Dungworth et McDonald (1980), il s'agit donc plutôt du degré de l'occurrence de certains marqueurs grammaticaux :

The nature of language is such that general language and special languages can be accommodated within one natural language: the fundamental characteristics of language are manifested both in English and in the language of chemical engineering, both in French and in the language of physics. The differences between general and special languages is a difference of degree rather than kind: the degree to which the fundamental characteristics of language are maximised or minimized in special languages. (Sager et al., 1980, p.38, cité par Pic et Furmaniak, 2010, p.2-3)

Ainsi, certains des marqueurs peuvent être plus fréquents, ou, à l'inverse, presque absents dans une langue de spécialité. Ils peuvent également changer de valeur comme c'est le cas des modaux en anglais juridique qui très souvent perdent leur valeur épistémique (Pic & Furmaniak, 2010).

Plusieurs chercheurs (Petitjean (1989), Rastier (2001), Krazem (2011), Adam (2011)) soulignent l'impossibilité de classification des genres à partir des caractères formels étant donné que ces caractères ne surviennent pas tous dans même temps, p.ex. un seul article scientifique ne comprend pas tous les caractères formels d'article vu comme un genre. C'est pourquoi il faudrait appuyer la classification sur « des groupements d'attributs d'importance variable » et pas sur « la recherche de critères définitoires en termes de conditions nécessaires et suffisantes » (Adam, 2011 : 117). La présence de certains schémas syntaxiques ne caractérise pas un genre d'une manière univoque ; le même schéma peut aussi apparaître dans des genres bien éloignés l'un de l'autre. C'est pourquoi il faut toujours s'interroger sur la fonction pragmatique de l'énoncé et ne pas interpréter le texte en se basant uniquement sur un schéma syntaxique apparu (Freyermuth S., p. 185). De fait, la présence de certaines constructions grammaticales peut dissimuler un autre genre. Ainsi, un texte romanesque peut prendre forme d'un genre procédural (une recette de cuisine ou une instruction) grâce au recours au mode injonctif et l'introduction de l'instance de l'allocutaire. Malgré ces ressemblances formelles, le genre est régi par d'autres critères et requiert une autre interprétation.

La fréquence élevée d'un fait de langue dans un texte donné n'est pas un facteur suffisant pour distinguer un genre (Beacco, 2004). À côté d'une consistance et d'une régularité de l'apparition du fait linguistique dans un genre, il faut prendre en compte sa représentativité, c'est-à-dire le caractère d'indicateur d'un genre. C'est ce que Beacco

appelle le caractère signalétique socialement ancré du marqueur linguistique d'un genre, telle la tournure *Il était une fois* qui désigne tout de suite un conte malgré sa seule occurrence dans le récit.

Un des linguistes contemporains saillants reliant la notion du genre et celle de la grammaire est Biber (2006). En se servant de grands corpus, il repère les faits de langue ayant une fonction communicative spécifique et étant sur- et sous-représentés. Biber a travaillé au sein du projet TOEFL 2000 Spoken and Written Academic Language (T2K-SWAL), projet qui devait répondre à la nécessité d'une norme externe pour évaluer la représentativité des matériaux ESL/EFL (English as a Second Language/English as a Foreign Language). La première partie du projet consistait en la construction d'un corpus qui comprenait les registres universitaires oraux et écrits, toutes les principales disciplines universitaires et tous les niveaux académiques. Dans son corpus, Biber a aussi inclus des registres « institutionnels », comme les manuels, les brochures, les pages web présentant les cursus, les groupes d'études, les rencontres avec le personnel. Ensuite, il a analysé les indices formels, ainsi que la distribution du vocabulaire et des ensembles lexicaux en prenant en compte les différences associées au registre, à la discipline et au niveau académique. La troisième étape du projet consistait à développer des outils de diagnostic qui analysent des caractéristiques linguistiques d'un texte et, par conséquent, ces outils permettent d'évaluer dans quelle mesure un texte est représentatif d'un registre ciblé. Ainsi, il établit une typologie de genres tels que les genres à haute densité d'information, les genres narratifs, les genres persuasifs.

Tout comme Bakhtine, nous pensons qu'il est indispensable de prendre en compte l'aspect grammatical lors de l'étude linguistique du genre. Nous soutenons que la fréquence de certains marqueurs linguistiques, ainsi que leur valeur, peut changer selon un genre étudié. De plus, il serait opportun qu'une étude de ce type advienne à l'aide du corpus comme l'a initié Biber pour le langage scientifique. Dans la section suivante, nous mentionnerons quelques traits du discours scientifique relevés dans des travaux précédents.

2. Caractéristique du discours scientifique selon les travaux précédents

Dans cette section, nous voudrions nous pencher sur les traits principaux de la langue scientifique relevés dans des travaux précédents en nous concentrant sur les niveaux morphologique et syntaxique. Nous nous appuierons principalement sur deux ouvrages qui traitent ce sujet : *La langue française de la technique et de la science* de R. Kocourek

(1982, désormais KOC), *University Language : A corpus based study of spoken and written registers* de D. Biber (2006, désormais BIB), ainsi que sur les travaux réalisés dans le cadre du projet Scientext : le travail sur le lexique verbal et positionnement de l’auteur dans les articles en sciences humaines (Tutin, 2010) et sur la phraséologie transdisciplinaire des écrits scientifiques (Tran, 2014). Bien entendu, on peut souligner des divergences entre ces recherches : Kocourek analyse la langue française et il se concentre uniquement sur l’écrit alors que Biber s’occupe de la langue anglaise avec une perspective plus large car il prend en compte la variation diamesique et il relève des divergences entre des disciplines différentes. De plus, les deux ouvrages se distinguent par un écart temporel assez important de plus de trente ans. Au plan méthodologique, Kocourek a parcouru les textes en les étudiant manuellement, leur nombre était restreint, tandis que Biber se base entièrement sur des méthodes informatisées basées sur des corpus sous format électronique. Malgré cela, en ayant recours à ces deux ouvrages, nous essayerons de capter quelques traits du langage scientifique que nous explorerons plus en détail par notre étude de corpus. Notre travail est en grande partie au départ intuitif et c’est grâce au corpus que nous vérifierons nos hypothèses. Ce répertoire nous permettra, donc, de faire un petit bilan de l’état des choses actuelles et poser des hypothèses pertinentes. Nous avons divisé les traits en trois groupes : aspects lexicologiques, aspects syntaxiques et grammaticaux, aspects textuels.

1) **Aspects lexicologiques**

a) **Sémantique des noms**

En ce qui concerne la sémantique des noms, les substantifs abstraits constituent la moitié de tous les noms tant dans des registres écrits que dans des registres oraux. (BIB) Parmi ceux-ci, les noms qui ont trait aux fonctions intellectuelles de l’esprit (« mental nouns » : *decision, experience, conclusion, expectation, observation, recognition, assumption* Biber, 2006, p.53) sont plus fréquents dans des disciplines liées aux sciences humaines. De plus, les sciences humaines ont recours à des noms hautement spécialisés. Dans les textes de gestion et dans les textes d’ingénierie, on retrouve plutôt des noms concernant des concepts et des processus abstraits. Les textes d’ingénierie se caractérisent aussi par un grand nombre de noms concrets et l’absence des noms animés. Cela reflète la spécificité de ce type des textes. De fait, ils décrivent des entités de la vie quotidienne en termes techniques.

b) Sémantique des verbes

Une des disciplines les plus productives quant aux verbes est la linguistique (BIB). Elle utilise particulièrement les verbes renvoyant aux opinions, aux intentions, à la démonstration, aux résultats et aux choix effectués. La psychologie, qui utilise beaucoup moins de verbes, utilise ceux qui renvoient à la dimension expérimentale : les verbes d'hypothèse, de résultats et d'opinion. Dans les sciences de l'éducation, ce sont des verbes d'intention, d'opinion, de question et de résultats qui sont les plus fréquents (Tutin, 2010). Ils permettent d'exprimer les motivations et les raisons d'être de la recherche.

2) Aspects syntaxiques et grammaticaux

a) Parties de discours

D'après Biber (2006), dans les registres écrits, la partie de discours la plus utilisée est celle des noms. (BIB) Les verbes ne transmettent pas beaucoup d'informations, car leur fonction primaire est de lier de longues et complexes propositions nominales alors que les registres parlés utilisent le même nombre de noms et de verbes. Les adjectifs sont utilisés le plus souvent dans les registres écrits, tandis que les adverbes sont favorisés dans les registres parlés. Dans les registres parlés, les noms sont plus simples et l'utilisation des pronoms est plus récurrent qu'à l'écrit.

b) Temps verbaux

En ce qui concerne la variation diachronique et l'emploi de temps grammaticaux, soit dans l'écrit soit dans des registres parlés, le passé est un temps marqué, ce que veut dire que c'est un temps peu usuel qui est employé dans des buts spécifiques. Les plus grandes différences quant à l'emploi du temps sont observées entre plusieurs disciplines scientifiques. À une extrémité, nous avons l'ingénierie avec un emploi très faible de temps du passé (5 % de tous les verbes) et à l'autre, il y a les sciences humaines où l'emploi du passé est relativement commun (40 % de tous les verbes dans des manuels universitaires). Cela peut être dû à la thématique et la spécificité de ce type de textes ; de fait, ils racontent très souvent les événements passés et, le développement des théories se passe à travers le parcours des idées des prédécesseurs.

KOC observe un phénomène similaire : c'est le présent qui prédomine dans des écrits scientifiques (85 % de toutes les formes verbales), 10 % des occurrences étant réservées au passé composé et au futur ; le reste est réparti entre les autres temps et modes. Le futur et le passé composé permettent de créer des liaisons intraphrastiques. Pour cela on trouve souvent le futur dans des introductions (p.ex. « Nous n'envisagerons ici que le pétrole *nouveau* à découvrir », Kocourek, 1982, p. 51) et le passé composé dans les conclusions (« Cette étude nous a permis de compléter les données physicochimiques disponibles. » Kocourek, 1982, p. 51).

Les autres chercheurs relèvent le même phénomène : Hoffmann (1976) constate que l'écrit scientifique perd son rapport temporel concret. Il y a également ceux qui soulignent son caractère aspectuel : « une perspective atemporelle » de Vigner et Martin (cité par Kocourek, 1982, p.51).

c) **Passif et actif**

Tout comme le présent, la voix active est un choix non marqué dans des écrits scientifiques. Tous les registres parlés utilisent les phrases verbales à la voix active plus de 95 % du temps (BIB). En revanche, les registres universitaires écrits démontrent un plus grand emploi de la voix passive : 20 % de toutes les phrases verbales. En ce qui concerne les disciplines, comme le démontre Biber, c'est l'ingénierie qui a le plus recours à la voix passive (presque 30 % de toutes les phrases verbales).

d) **Pronoms**

Dans la communication écrite, le système des pronoms est assez restreint. (KOC) Les pronoms de la deuxième personne *tu* et *vous* sont quasiment absents. L'auteur, lui, n'emploie que très rarement *je* (réservé plutôt aux textes moins formels et aux écrits-parlés et utilisé en linguistique, Flottum et al., 2006) qui est remplacé par *nous*, soi-disant *nous de modestie*. Pourtant, *nous* dans l'écrit scientifique peut avoir aussi les autres emplois :

- L'emploi traditionnel. Vu que de nos jours il y a beaucoup de recherches collectives, *nous* peut se référer aux plusieurs auteurs

- *Le nous inclusif*. Le *nous* dans des textes scientifiques a tendance à inclure, en plus de l'auteur ou des auteurs, le destinataire latent (p.ex. « Notons que les

termes (...) doivent être considérés comme des potentiels effectifs » (Kocourek, 1982, p.50) ; dans un autre type de texte *notons* serait plutôt remplacé par *notez*).

Le pronom *on*, qui désigne des êtres humains d'une façon indéfinie et impersonnelle, joue un rôle important dans des textes écrits scientifiques. Tout comme *nous*, le pronom *on* représente l'auteur ou des auteurs, avec ou sans inclusion du destinataire. Tutin (2010) a démontré le lien entre l'emploi des pronoms personnels et la prise de position par l'auteur : plus les verbes expriment une position marquée, moins ils sont assumés par l'auteur. Ainsi, les verbes d'opinion et d'évaluation sont le plus souvent introduits par les pronoms *on* et *nous* inclusifs, pendant que les pronoms exclusifs sont plutôt utilisés avec les verbes indiquant un apport scientifique ou une intention. Par ailleurs, les pronoms inclusifs accompagnent les verbes à fonction évidentielle qui démontrent le déroulement de la recherche, comme le verbe *voir* et ont une fonction dialogique forte.

En dehors de *nous*, *on* et le pronom invariable *il*, les textes scientifiques ne contiennent que les pronoms des 3 personnes du singulier et du pluriel anaphoriques : *il*, *le*, *lui*, *eux*, *leur*, *elle(s)*. Cependant, ils ne sont pas très fréquents et ils fonctionnent comme des éléments de cohérence textuelle ou ils sont un lien intraphrastique entre deux propositions de la même phrase.

e) **Structures complexes**

Tout d'abord, un fait remarquable est la longueur de la phrase dans l'écrit scientifique. (KOC) Dans la langue française, la longueur moyenne de la phrase dans le langage scientifique se place entre 28-29 mots. Avec les mêmes résultats, nous pouvons trouver en anglais (29 mots dans un ouvrage scientifique contre 8 mots des bandes dessinées) et en russe (28,5 mots dans la prose scientifique contre 8 dans des romans esthétiques). Pour voir encore mieux la complexité de la phrase indépendante, nous pouvons examiner une phrase indépendante déclarative à 30 mots:

« La comparaison des cartes de distribution des prises de bonites à celle de la salinité de surface montre des relations très nettes entre l'importance des prises et la salinité » (Kocourek, 1982, p.53)

Nous voyons que la complexité de la phrase résulte des syntagmes nominaux qui peuvent remplir diverses fonctions : sujet, complément d'objet, attribut, circonstant,

modificateur. La phrase citée est constituée de deux syntagmes nominaux principaux (avec les noyaux : *comparaison* et *relations*) qui peuvent être analysés en des syntagmes moins complexes. De fait, Kocourek observe pour la langue française la même chose que Biber pour la langue anglaise : une forte présence des nominalisations dans les phrases des écrits scientifiques.

Selon Biber, contrairement à ce qu'on pourrait supposer, les subordonnées, associées à une certaine complexité linguistique, sont généralement plus utilisées dans des registres parlés que dans les registres écrits. Si on étudie la subordonnée du langage scientifique d'encore plus près, on voit que la subordonnée adverbiale et complétive prévaut dans des registres parlés, alors que la subordonnée relative est plus fréquente dans des registres écrits.

Un autre trait syntaxique est la prédominance du type déclaratif sur les autres types des phrases (KOC). Les phrases interrogatives sont principalement restreintes à deux emplois possibles:

- La phrase oratoire qui, en renvoyant à la phrase suivante, contribue à la cohésion du texte p.ex. *Comment interpréter les inversions du rendement ?* (p.52, KOC)
- Une incise interrogative exprimant une explication possible suggérée. p.ex. *Depuis, les effectifs sont inférieurs à ce chiffre malgré une excellente prospection (peut-être 1971 fut-elle une excellente année de reproduction pour cette espèce ?)* (ibidem, p.52)

En ce qui concerne l'impératif, il est principalement utilisé à la première personne du pluriel (*rappelons, ajoutons, supposons, mettons, posons, remarquons*) dans l'emploi de *nous inclusif* pour renforcer le lien avec le destinataire latent.

3) Aspects textuels - cohérence textuelle

Selon Kocourek, le type de cohérence le plus fréquent de l'écrit scientifique est la reprise des liens sémantiques des phrases voisines. Cela se fait très souvent en recourant à la répétition partielle d'un élément (grâce aux articles définis ou aux démonstratifs *ce, cette, ces*) ou à son remplacement complet (remplacement du nom spécifique par son hyperonyme, toujours en employant le démonstratif ou l'article défini).

La cohérence interphrastique peut être également d'ordre grammatical quand on a recours aux pronoms, déterminants, conjonctions. Le plus récurrent, pourtant est l'emploi anaphorique des pronoms et des déterminants qui remplacent un syntagme d'un syntagme précédent, p.ex. :

« **Les matériaux composites à fibres** sont souvent utilisés du fait que, pour la même résistance ou le même module, **ils** sont nettement plus légers que les matériaux traditionnels, par exemple les métaux. C'est-à-dire que **leurs** propriétés spécifiques sont souvent d'importance majeure (...) » (Kocourek, 1982, p.40).

Kocourek mentionne aussi des unités de liaisons (conjonctions, adverbes, locutions) qui aident à maintenir la cohérence intraphrastique. Les unités de liaisons de Kocourek sont les connecteurs de Biber. Selon lui, ils servent à lier des phrases, les parties du discours entre eux, et ils en indiquent les relations logiques et fournissent un cadre interprétatif pour l'auditeur/le lecteur. Les connecteurs peuvent être divisés en marqueurs du discours et adverbes de liaison. Les premiers (p.ex. *well, so, ok*) sont utilisés dans des registres parlés. Les adverbes de liaison peuvent être considérés comme un système parallèle au discours marqueurs qui, cette fois, fonctionnent plutôt dans des registres écrits. À ce point, nous voudrions mentionner le travail de Tran (2014). Dans sa recherche, en se limitant aux unités polylexicales, elle a distingué les marqueurs métadiscursifs et les marqueurs logiques. Les premiers ont un rôle d'organisation textuelle participant à la structure globale du texte, les deuxièmes sont des marqueurs à fonction argumentative qui contribuent à établir le lien interne du texte.

La consultation des travaux précédents sur le langage scientifique nous a montré certaines de ses spécificités : la complexité de la syntaxe, l'impersonnalité de la phrase, la passivation fréquente, l'emploi restreint des pronoms, l'usage courant des adverbes et des unités polylexicales assurant la cohérence textuelle. Ces observations nous dirigent vers des constructions potentielles qui semblent largement représentées dans l'écrit scientifique telles que les constructions impersonnelles, passives, infinitivales, circonstancielles et dont nous voudrions véritablement vérifier la surreprésentation à l'épreuve des corpus dans des écrits scientifiques.

3. Constructions syntaxiques

Pour notre recherche, qui consiste à repérer et à décrire les constructions surreprésentées dans l'écrit scientifique, nous allons nous pencher sur les courants syntaxiques qui, d'une certaine façon, relient la syntaxe et la sémantique; nous partons de

l'idée qu'une construction grammaticale, au-delà de la forme, contribue au sens transmis par le message. Nous choisissons essentiellement deux modèles qui s'inscrivent dans ce courant : le lexique-grammaire et les grammaires de construction.

A. Le lexique-grammaire

Le lexique-grammaire, dont les idées principales ont été formulées pendant les années 70 par M. Gross (1975), est une approche syntactico-sémantique. Elle se base sur l'hypothèse que les relations entre les diverses structures syntaxiques subissent de fortes contraintes lexicales. La description syntaxique devrait donc contenir, à côté de règles générales, des précisions de type lexical : quel mot exige, autorise ou interdit l'emploi de quelle règle. (Gardent C. et al., 2005).

La description consiste à repérer toutes les constructions et les distributions des items lexicaux concernés : verbes, noms prédictifs et adjectifs. Plus précisément, pour donner un exemple, il y a une classification dans des groupes de tables de tous les verbes simples du français selon leurs propriétés syntaxiques et distributionnelles. Les lignes de tables correspondent aux items lexicaux et les colonnes aux propriétés syntaxiques ou sémantiques; à chaque croisée il y a un signe « + » si le verbe possède la propriété ou « – » dans le cas contraire. S'il existe des doutes concernant l'acceptabilité d'une construction, elle est considérée comme correcte:

Les exemples cruciaux étant de manière générale les séquences inacceptables, une première tactique a été, en cas d'hésitation sur l'entrée d'un élément dans une structure et d'absence d'une argumentation de niveau supérieur à celui du jugement empirique ponctuel, de considérer comme mal formée une structure qui ne l'est pas est pour nous plus grave que le risque inverse, inacceptable pour le puriste, d'accepter une structure « non recommandable ». (cité par Yannick Mathieu, 2009, p. 8).

En ce qui concerne la méthode de jugement sur l'acceptabilité ou non d'une construction c'est majoritairement l'introspection et dans un second temps, la consultation des autres travaux descriptifs et des dictionnaires qui ont été employées.

Le français est un français standard, révélé essentiellement par introspection, et vérifié, pour certains exemples douteux, par enquête auprès de locuteurs d'appartenance sociolinguistique voisine; les grammaires et ouvrages de linguistique traditionnels ainsi que les dictionnaires nous ont également fourni de nombreux exemples. (cité par Yannick Mathieu, 2009, p. 9).

Dans ce modèle, il n'existe pas de règles grammaticales entièrement indépendantes; elles sont toujours soumises aux particularités lexicales des unités

auxquelles elles sont applicables. (Ibrahim A., 2010). De plus, les mécanismes formels de la grammaire, qui créent une prédication, ont eux aussi une valeur sémantique.

Pendant les premiers vingt ans, le travail sur le lexique-grammaire était plutôt descriptif pour pouvoir être exploité dans la réalisation effective, voire pour la conception de structures de données et d'algorithmes pour l'analyse syntaxique ce qui advient dès le début des années 1990. Finalement, pour la langue française, on compte 15 000 entrées de verbes simples et, si on compte les autres fragments (noms prédictifs, phrases figées, adverbes figés), le nombre s'élève à 15 000 entrées. (Laporte, E. 2005).

Cette approche n'est pas basée sur corpus, contrairement à la nôtre, mais le classement du lexique-grammaire démontre la structure des constructions les plus significatives, leurs propriétés distributionnelles et transformationnelles. Cependant, les tables du lexique-grammaire ne sont pas exhaustives et elles ne prennent pas en compte les fonctions discursive ou rhétorique.

B. Les patrons verbaux de l'écrit scientifique – recherche de Yan (2013)

À ce point, après avoir signalé deux courants reliant l'aspect sémantique et grammatical, nous voudrions mentionner la recherche de Yan (2013). Elle a relevé les patrons verbaux de vingt verbes « pleins » les plus fréquents appartenant au « lexique transdisciplinaire des écrits scientifiques » (Tutin, 2007), à savoir « le lexique partagé par la communauté scientifique mis en œuvre dans la description et la présentation de l'activité scientifique » (ibid. : 190) dans le but de concevoir des outils didactiques appropriés à la didactique du français de spécialité ou français sur objectifs spécifiques (FOS). Yan a repéré des lemmes verbaux et les patrons lexicosyntaxiques dont le lemme verbal fait partie. Tout comme les constructions du lexique-grammaire, chaque patron basé sur un lemme verbal est associé à un sens spécifique et se compose des structures syntagmatiques transmettant des valeurs sémantiques.

Yan s'inspire du modèle de patron de Hanks (2008) élaboré dans le cadre du projet the Corpus Pattern Analysis (CPA, analyse des patrons basée sur le corpus). L'approche de Hanks consiste en l'identification d'un verbe individuel et du repérage de ses patrons et des sens qui en résultent. Cette approche relève donc à la fois de la syntaxe, mais elle met aussi l'accent sur la sémantique. Ainsi, le modèle de Hanks permet de désambiguïser les structures qui ayant la même structure syntaxique, présentent des sens différents, p.ex. Fire [[Projectile]] from [[Firearm]] vs Fire [[Human]] from [[Activity]].

Pour sa recherche Yan a défini le patron verbal comme:

une construction autour du verbe et dont ce verbe fait partie ; cette construction composée de plus de deux mots est associée à un sens clair du verbe ; chaque élément du patron verbal se forme sous une structure syntagmatique ; chaque valence ayant une valeur sémantique comprend des ensembles lexicaux de noms ou d'autres mots qui sont souvent des synonymes. (Yan, 2013, p.23)

Yan a limité sa recherche aux patrons verbaux des 20 verbes « pleins » (sans les verbes auxiliaires, les verbes modaux, les verbes support) les plus fréquents dans l'écrit scientifique. Son outil principal de repérage des patrons était le corpus Scientext, grâce auquel, elle a examiné des contextes dans les concordances de ces 20 verbes. Après une première observation, Yan a repéré les structures syntaxiques caractéristiques pour le verbe recherché (il fallait plus de vingt occurrences d'une structure pour que le verbe soit repéré). Une fois les structures syntaxiques dégagées, on passait à leur examen : on discernait les propriétés sémantiques des éléments constitutifs:

1. [[Personne]] sait (ADV [[Quantité| logique| temps]]) que...
2. [[Personne]] le sait (ADV [[Quantité| logique| temps]]) (dans des phrases à incise)
3. [[Personne]] ne sait pas [Proposition introduite par si]
4. [[Evènement |Entité abstraite]] [[ne]] saurait/sauraient (ADV [[Logique]]) [V-inf.| être + participe passé]
5. [[Personne]] [[ne]] saurait/sauraient (ADV [[Logique| affirmation/doute]]) [V-inf.]

Nous voudrions signaler les ressemblances de l'étude de Yan avec la nôtre puisque nous nous concentrons aussi sur l'aspect syntaxico-sémantique des constructions. La différence consiste dans l'approche : Yan part des lexèmes concrets et elle recherche ensuite des constructions syntaxiques plus générales d'un lexème en repérant en même temps certaines classes sémantiques de ses constituants ; nous partons des constructions générales et seulement alors, nous repérons les lexèmes qui font partie de cette construction avec les fonctions discursives et pragmatiques des constructions.

C. Grammaires de constructions

Un modèle plus récent associant la forme et les sens est celui de la famille des grammaires de constructions qui recouvrent l'ensemble des modèles et de théories qui assument comme une unité fondamentale du langage la construction, c'est-à-dire une entité qui associe strictement le sens et la forme. Les modèles grammaticaux des grammaires de constructions tirent leurs origines des propositions de la linguistique cognitive. (Pukicka, 2008). Parmi les représentants les plus éminents dès leur début jusqu'à des temps plus récents, on peut mentionner : R. Langacker (*Cognitive Grammar* 1987, 1991, 1997), Ch. Fillmore et P. Kay (*Construction Grammar*, 1988, 1993) A. Goldberg (*Cognitive Construction Grammar*, 1995, 2006) et W. Croft (*Radical Construction Grammar* (2001). Desagulier G. (2007) voit dans les grammaires de construction une réaction des cognitivistes aux théories générativistes de la syntaxe et, en même temps, le retour aux théories linguistiques de Saussure ; de fait, tous les deux convergent sur la nature mentale et conventionnelle du signe linguistique. Dans les deux approches, l'image graphique/acoustique du signe et le concept sont présents uniquement dans l'esprit du locuteur, le signe n'ayant pas de référent objectif dans la réalité.

Selon Mathieu (2003), une construction, qui est une notion purement théorique, permet de créer des « constructs », c'est-à-dire des éléments de la langue : les mots, les syntagmes, les phrases. Elle rapproche les GCx de la grammaire générative en tant qu' « un ensemble de constructions hiérarchisées (y compris les mots) qui peuvent être unifiées pour créer l'ensemble des phrases acceptables d'une langue, mais qui ne génèrent aucune phrase non acceptable. » (Mathieu, 2003, p. 43)

Desagulier (2007) en se référant à la conception du signe saussurien, définit une « construction » comme « un symbole intrinsèquement porteur de sens » (Desagulier, 2007, p. 1). Pour lui, c'est une unité de base. En tenant en compte les propos d'un des représentants des GCx, Goldberg, il rappelle qu'une construction est une « entité indépendante » (cité par Desagulier, 2007, p. 1).

En s'appuyant sur la théorie de G. Lakoff, François (2008) définit quant à lui les constructions comme des paires conventionnalisées forme-sens qui « constituent un inventaire structuré et ils sont liées mutuellement par des généralisations systématiques, des relations d'héritage et des liens de catégorisations. » (François, 2008, p.8). Une construction peut être valide, si elle satisfait trois conditions:

- a. Sa forme est corrélée directement avec un sens ou une fonction déterminée

- b. Sa forme ne se laisse pas (complètement) dériver d'autres formes ;
- c. Sa sémantique n'est pas (complètement) compositionnelle.

Il cite la définition de Croft & Cruse (2004) :

une grammaire de construction consiste en un grand nombre de constructions de tous types depuis les constructions syntaxiques schématiques jusqu'aux items lexicaux substantifs. Toutes ces constructions possèdent des propriétés de forme (syntaxiques et phonologiques) et de sens (sémantiques et pragmatiques). Toutes ces constructions sont organisées de manière particulière dans l'esprit du locuteur. (cité par François, 2008, p.8)

Selon Mathieu (2003), un objectif principal de grammaires de construction est de décrire entièrement une langue en tenant compte des niveaux syntaxique, phonologique, morphologique, sémantique et pragmatique. Elle souligne le fait que les expressions idiomatiques, expressions figées, etc. ont le même statut que les constructions « normales » : elles ne sont pas aux marges de la langue comme dans l'approche traditionnelle. Outre cela, selon Mathieu (2003), la GC n'a pas les caractéristiques de grammaire transformationnelle. Ainsi, la forme passive n'est pas considérée comme une dérivation de la forme active, mais est une forme à part, indépendante. Une des autres particularités de GC est le fait qu'une diversité entre des formes syntaxiques a toujours ses implications sémantiques ou pragmatiques. De plus, la somme de composants d'une construction ne donne pas toujours un sens prédictible. À ce sujet, François (2008) ajoute que critère de compositionnalité varie selon les auteurs ; écarté par Langacker il est requis par Goldberg. Selon Goldberg, Fillmore, Croft les constructions existent indépendamment l'une à côté de l'autre, alors que Langacker soutient que les constructions « abstraites » dérivent des constructions plus spécifiques par le processus d'abstraction.

Dans le cadre des grammaires cognitives, nous voudrions introduire le projet The Swedish Construction (Lyngfelt et al., 2012), basé sur FrameNet qui, à son tour, est réalisé à l'International Computer Science Institute à Berkeley en Californie. Le projet s'appuie sur la Sémantique des frames de Fillmore (Fillmore, 1976), frames étant définis comme structures cognitives associées au processus de compréhension. La thèse principale de la Sémantique de frames est que nous ne serions pas capables de comprendre un énoncé sans les connaissances indispensables relatives à cet énoncé, p.ex. pour comprendre le verbe *vendre*, il est essentiel de connaître la situation de transfert commercial, à savoir la relation entre le vendeur et l'acheteur, la relation entre l'argent et les marchandises, les relations entre le vendeur et les biens et ainsi de suite. Une des caractéristiques du projet The

Swedish Construction est la prise en compte du scénario d'une construction donnée dans ce cadre.

La description d'une construction dans le cadre de The Swedish Construction comporte la définition d'une construction en texte libre, les descriptions structurelles schématiques, les définitions des éléments d'une construction et des exemples annotés :

reflexiv_resultativ

type	Cx
category	vbm
evokes	Causation_scenario
definition	[Någon] _{Actor} eller [något] _{Theme} utför eller undergår [en aktion] _{Activity} som leder (eller antas leda) till att [aktören] _{Actor} / [temat] _{Theme} , uttryckt med reflexiv, uppnår [ett tillstånd] _{Result} .
structure	vb refl AP
cee	refl
coll	{äta ¹ : mätt ¹ } {supa ¹ : full ² } {skrika ¹ : hes ¹ } springa ¹
internal construction elements	<ul style="list-style-type: none"> ▪ role: name=Activity cat=vb ▪ role: cx=refl name=Actor ▪ role: cx=refl name=Theme ▪ role: name=Result cat=AP
external construction elements	<ul style="list-style-type: none"> ▪ role: name=Actor cat=NP ▪ role: name=Theme cat=NP
examples	<ul style="list-style-type: none"> ▪ [Vi åskådare]_{Actor} [[springer]_{Activity} [oss]_{Actor} inte [varma]_{Result}]resultativ_reflexiv direkt. ▪ [Kornet och havren]_{Theme} får [[frysa]_{Activity} [sig]_{Theme} [mogen]_{Result}]resultativ_reflexiv. ▪ [[Drick]_{Activity} [dig]_{Actor} [smal]_{Result}]resultativ_reflexiv i vår.
comment	Det finns också en PP-variant med resultatativ betydelse, t.ex. "träna sig i form", som ev. bör inkorporeras här - alt. betraktas som en metaforisk utvidgning av någon rörelse-cx.
reference	Jansson, Håkan (2006): Har du ölat dig odödlig? En undersökning av resultativkonstruktioner i svenskan. (D-uppsats, även publicerad som MISS 57) http://hdl.handle.net/2077/19000 Lyngfelt, Benjamin (2007): Mellan polerna. Reflexiv- och deponenskonstruktioner i svenskan. Språk och stil NF 17: 86–134. http://hdl.handle.net/2077/21731

Un exemple d'une construction de The Swedish Construction

La construction est décrite en commençant par le type (Cx) et la catégorie (vbm) à laquelle elle appartient. Ensuite, on mentionne le scénario qui indique la situation communicationnelle. La définition caractérise la construction en attribuant les rôles sémantiques à ses constituants ; ces rôles sont encore plus clairement indiqués grâce à la division en éléments de la construction internes et externes. Dans notre exemple, les rôles concernent les participants de l'activité ou les thèmes (dans le cas des noms non animés),

l'activité et le résultat de l'activité. À côté de la valeur sémantique, on indique les éléments grammaticaux ; la structure est également indiquée en termes purement grammaticaux. Par la suite, la description contient des exemples authentiques, le commentaire supplémentaire et les références aux matériaux utilisés lors de la description.

Nous nous inspirons de ce modèle pour décrire nos constructions, car il relie l'aspect grammatical avec la sémantique tout en restant clair et compréhensible. Cependant, comme nous l'indiquerons dans le chapitre suivant, nous voulions simplifier ce modèle. Nous n'utilisons pas de rôles sémantiques et pas de métalangage formel.

Les ressources lexicales concernent plutôt les mots, tandis que les grammaires se concentrent sur l'aspect morphologique ou syntaxique. Notre approche nous a amenée à relier l'aspect sémantique et l'aspect grammatical des constructions analysées. Pour la définition de la construction, nous nous appuyons sur celle fournie par Lyngfelt et al. (2012), travaillant pour le projet The Swedish Construction, selon qui :

Language consists to a quite large extent of semigeneral linguistic patterns, neither general rules of grammar nor lexically specific idiosyncracies. Such patterns may be called constructions.

La construction est une structure partiellement schématique, elle comprend donc des éléments stables et changeants. En nous basant sur cette définition, ainsi que sur la description de The Swedish Construction, nous essayerons d'établir notre propre modèle de description.

4. Le modèle de construction adopté

Dans ce travail, nous nous sommes inspirée du modèle de description des constructions du projet The Swedish Construction. Cependant, nous avons renoncé aux rôles sémantiques et au métalangage formel. Notre objet d'étude est le langage scientifique, et nous voulons donc aussi insister sur la dimension rhétorique.

Pour donner un exemple de description de construction, nous prenons une première construction que nous avons étudiée : **II + être + ADJ + que + ind./subj.** Pour aborder l'analyse des constructions, nous allons donc élaborer un modèle de description :

Tout d’abord, nous indiquons la structure, à savoir les composants grammaticaux d’une construction. Ensuite, dans la définition, nous élargissons cette description en ajoutant des fonctions logiques des composants. La partie suivante concerne les propriétés morphologiques : nous indiquons les propriétés stables des constituants de la construction. La tête lexicale ce sont des éléments essentiels de la construction ; ce sont les lexèmes qui apparaissent dans la construction dans les articles du corpus Scientext. Ils nous ont permis de distinguer les fonctions de la construction. Ensuite, nous donnons des exemples puisés dans le corpus. Dans le commentaire, nous complétons les informations, en précisons la fonction rhétorique ou discursive de la construction et en indiquant le temps et le mode de certains de ces constituants (ce qui très souvent, influence la sémantique de toute une construction).

Dans le prochain chapitre, nous décrirons plus précisément les étapes méthodologiques qui nous ont permis à identifier les constructions, ainsi que l’outil de traitement de corpus utilisé, à savoir le corpus Scientext.

La structure	Il + V + ADJ + que + ind./subj.
La définition	<p>Il [sujet apparent] + verbe copule + adjectif [attribut du sujet] + que + proposition [indicatif ou subjonctif]</p> <p>Une construction impersonnelle à sujet phrastique extraposé, composée du pronom impersonnel <i>il</i> (appelé également le sujet pléonastique ou explétif) suivi d’un verbe copule, un adjectif et la conjonction de subordination <i>que</i> qui introduit le complément phrastique.</p>
Les constituants et leurs propriétés morphologiques	<p>Il – pronom impersonnel, m, sg</p> <p>Verbe, 3 per, sg, indicatif</p> <p>Adjectif, m, sg</p> <p>Que – conjonction de subordination</p>
La tête lexicale	<p>ADJ : étonnant, évident, absurde, acceptable, banal, bon, clair, commode, commun, difficile, erroné, excessif, fréquent, fructueux, important, impossible, indéniable, indispensable, inintéressant,</p>

	<p>intéressant, inutile, légitime, mauvais, nécessaire, optimal, piquant, possible, probable, radical, rare, remarquable, sûr, significatif, souhaitable, théorique, utile, vrai, vraisemblable</p> <p>VERB : être, sembler, rester, apparaître, devenir, demeurer, paraître</p>
LES FONCTIONS	
Modalité épistémique : évidence (vrai, clair, remarquable, certain, évident, indéniable, indispensable, sûr)	
Exemples	
<p>Il est vrai que la rareté des exemples et les mises en garde de l'auteur sur le caractère non confessionnel de la majorité de ces conflits empêchent d'énoncer des conclusions hâtives. [his-art-135-body]</p> <p>Il est en effet clair que ce ne sont pas toutes les données qui vont être pertinentes à chaque fois. [lin-art-38-body]</p> <p>Cependant, il est évident que les électeurs d'origine étrangère votent davantage pour des politiciens d'origine étrangère que ne le font la majorité des électeurs. [spo-art-153-body]</p>	<p>La construction remplit une fonction rhétorique ou discursive : elle sert à donner de la valeur axiomatique aux thèses de l'auteur de l'article ou elle permet d'introduire un nouveau sujet. Elle permet également de présenter les résultats de la recherche : c'est pourquoi sa fréquence relative est la plus élevée dans la conclusion.</p> <p>Le verbe est à l'indicatif présent.</p> <p>Le complément phrastique est à l'indicatif.</p>

Partie 2

-

Méthodologie

Chapitre 3. La recherche des constructions

Comme nous l'avons déjà mentionné dans la partie précédente, la plupart des travaux sur l'écrit scientifique se concentrent sur les aspects lexicaux. Dans le cadre de notre projet, nous avons essayé de repérer des constructions syntaxiques surreprésentées dans l'écrit scientifique. En ce qui concerne l'aspect heuristique de notre recherche, après avoir lu des travaux précédents sur le discours scientifique et, particulièrement, le travail de Kocourek, nous avons lu des articles scientifiques dans le but d'identifier des constructions potentielles. Nous avons listé des constructions susceptibles d'être surreprésentées dans l'article scientifique. Comme outil vérificatif par la suite, nous avons eu recours à deux corpus : un corpus d'écrits scientifiques et un corpus de contraste, à savoir un corpus des textes journalistiques et littéraires. Le premier est le corpus Scientext et le deuxième un corpus de presse disponible sur le site Scienquest.

1. Scientext

Le corpus Scientext comprend plusieurs corpus : des textes scientifiques en français, les textes scientifiques en anglais, des écrits universitaires d'étudiants d'anglais langue étrangère et des évaluations de communications. Nous nous sommes limitée à un sous-ensemble du premier corpus, et, plus précisément au corpus élaboré par Tran (2014). Il est composé d'articles appartenant à dix disciplines en SHS (linguistique, psychologie, sciences de l'éducation, économie, sciences politiques, anthropologie, histoire, géographie, sciences de l'information et de la communication, sociologie). Chaque discipline comporte 30 articles, sauf la psychologie et les sciences du langage qui contiennent 25 articles. La constitution du corpus a été basée sur le classement des revues, à savoir le classement de l'AERES (l'Agence de l'évaluation de la recherche et de l'enseignement supérieur). Sa classification comporte les notes A, B, C. Étant donné que c'est la note A qui assure une bonne qualité des articles publiés, seules les publications de cette catégorie ont été choisies. Pour assurer une certaine homogénéité du corpus, chaque discipline est représentée par deux à trois revues. Comme une revue est consacrée à une question spécifique des sous-domaines de chaque discipline, les articles proviennent des numéros différents et des auteurs différents pour différencier la thématique.

Le corpus Scientext a été annoté structurellement avec le standard XML TEI Lite, ce qui a permis de reconnaître les parties textuelles (introduction, conclusions, notes de bas

de page) et a été analysé au niveau de l'étiquetage morpho-syntaxique à l'aide de Syntex. Ainsi, chaque forme comporte un lemme, une partie de discours, des traits morphosyntaxiques (nombre, genre, temps, mode) et est reliée à d'autres mots par des dépendances syntaxiques. Ce trait est pour nous particulièrement important : nous pouvons rechercher des constructions plus générales, en nous basant aussi sur leur aspect grammatical est pas seulement lexical.

2. Le choix des corpus

Le corpus Scientext comprend plusieurs disciplines appartenant aux sciences humaines, sciences expérimentales et sciences appliquées. Dans notre recherche, nous nous limitons aux sciences humaines qui, en l'occurrence, comprennent : linguistique, psychologie, sciences de l'éducation, traitement automatique des langues, économie, sciences politiques, anthropologie, histoire, géographie, sciences de l'information et de la communication et sociologie ce qui donne 3 265 894 mots (357 textes). Comme type de documents, nous choisissons l'article avec toutes ses parties : introduction, développement, conclusions et résumé, notes, titres, remerciements, annexe.

1. Choix des types de textes Affiner la sélection: [Suite](#)

Ici, vous pouvez sélectionner un sous-ensemble du corpus, en fonction des disciplines, des genres et des parties textuelles.

357 textes sont sélectionnés (3 265 894 mots) [\[Voir les détails\]](#)

Disciplines	Types de documents	Parties
<input checked="" type="checkbox"/> Sciences humaines	<input checked="" type="checkbox"/> Article	<input checked="" type="checkbox"/> Parties principales
<input checked="" type="checkbox"/> Linguistique	<input type="checkbox"/> Communication	<input checked="" type="checkbox"/> Développement
<input checked="" type="checkbox"/> Psychologie	<input type="checkbox"/> Thèse	<input checked="" type="checkbox"/> Introduction
<input checked="" type="checkbox"/> Sciences de l'éducation	<input type="checkbox"/> HDR	<input checked="" type="checkbox"/> Conclusion
<input checked="" type="checkbox"/> Traitement Automatique des Langues		<input checked="" type="checkbox"/> Autres parties
<input checked="" type="checkbox"/> Économie		<input checked="" type="checkbox"/> Résumé
<input checked="" type="checkbox"/> Sciences politiques		<input checked="" type="checkbox"/> Notes
<input checked="" type="checkbox"/> Anthropologie		<input checked="" type="checkbox"/> Titres
<input checked="" type="checkbox"/> Histoire		<input checked="" type="checkbox"/> Remerciements
<input checked="" type="checkbox"/> Géographie		<input checked="" type="checkbox"/> Annexe
<input checked="" type="checkbox"/> Sciences de l'information et de la communication		
<input checked="" type="checkbox"/> Sociologie		
<input type="checkbox"/> Sciences expérimentales		
<input type="checkbox"/> Biologie		
<input type="checkbox"/> Médecine		
<input type="checkbox"/> Sciences appliquées		
<input type="checkbox"/> Électronique		
<input type="checkbox"/> Mécanique		

Tout Rien Tout Rien Tout Rien

Sauvegarder la sélection: [TEXTS] | Restaurer une sélection: Przeglądaj... Nie wybrano pliku. OK

Figure 1 : Choix de corpus

3. Le traitement du corpus

Sur Scientext, nous vérifions nos hypothèses grâce aux requêtes. Scientext permet trois types de recherche: la recherche sémantique (recherche à partir des grammaires créées

par l'équipe Scientext), la recherche libre et la recherche avancée. Nous démarrons en utilisant la recherche libre:

4. Recherche

Ici, vous pouvez chercher des occurrences dans le corpus.

Ou bien: perfectionner la requête en **Mode avancé**

Mots:

Mot 1: Forme Mot 2: Lemme Mot 3: Catégorie Mot 4: Lemme

Relation syntaxiques:

Relation 1: Mot 1 Mot 2

Relation 2: Mot 3 Mot 2

Attention, l'ordre des mots n'est plus pris en compte

Ajouter une relation

Recherche Interrompt arbitrairement la recherche à environ occurrences.

Figure 2 : Recherche libre sur le site Scientext

La recherche libre nous permet de choisir les forme, lemme et catégorie ; elle nous permet également d'établir les relations syntaxiques entre des mots particuliers. Dans le cas que nous présentons, il s'agit d'une construction impersonnelle : *il est ADJ que*. Nous choisissons donc respectivement comme forme *il* ou *Il*, le lemme, à savoir le verbe *être*, la catégorie d'adjectif et le lemme dernier c'est-à-dire la conjonction *que*. Ensuite, nous établissons les relations syntaxiques ; le premier mot (*il* ou *Il*) est le sujet du deuxième mot (le verbe *être*), le troisième mot (l'adjectif) fonctionne comme l'attribut du sujet du deuxième mot. De cette manière, nous obtenons 813 résultats si nous appuyons directement la touche *recherche*. Pourtant, tous les résultats ne sont pas pertinents : nous voyons beaucoup de résultats où l'ordre de mots n'est pas celui souhaité, telle la construction *qu'il est étranger*. Pour affiner notre recherche, nous récupérons la liste de lemmes adjectivaux apparaissant dans cette construction. Ainsi, nous obtenons tous les lemmes qui font partie de la construction précise *il est ADJ que* dans notre corpus. Ainsi, nous pouvons perfectionner notre recherche grâce au mode de recherche avancée.

4. Recherche

Ici, vous pouvez chercher des occurrences dans le corpus.

Éditez votre grammaire ci-dessous [\[documentation\]](#) :

```
$adje = étonnant, évident, absurde, acceptable, banal, bon, clair, commode, commun, difficile, erroné, excessif, fréquent, fructueux, important, impossible, indéniable, indispensable, intéressant, inutile, légitime, mauvais, nécessaire, optimal, piquant, possible, probable, radical, rare, remarquable, sûr, significatif, souhaitable, théorique, utile, vrai, vraisemblable
Main = <form=/^ (il|Il) $/, #1> %< <lemma=être, #2> %< <lemma=adje, #3> %< <lemma=que, #4> :: (SUJ, #1, #2) (ATTS, #3, #2);
```

Figure 3 : Recherche avancée Scientext

À partir de la liste de lemmes, nous créons une variable (\$adje). Ensuite, nous l'incluons dans notre requête au lieu de la catégorie générale de l'adjectif:

\$adje = étonnant, évident, absurde, acceptable, banal, bon, clair, commode, commun, difficile, erroné, excessif, fréquent, fructueux, important, impossible, indéniable, indispensable, inintéressant, intéressant, inutile, légitime, mauvais, nécessaire, optimal, piquant, possible, probable, radical, rare, remarquable, sûr, significatif, souhaitable, théorique, utile, vrai, vraisemblable

Main = <form=/^(il|II)\$/,#1> && <lemma=être,#2> && <lemma=adje,#3> && <lemma=que,#4> :: (SUJ,#1,#2)(ATTS,#3,#2);

Cette requête nous donne 576 résultats. Ensuite, nous procédons au traitement manuel du corpus : nous rejetons des résultats qui ne sont toujours pas pertinents en ce qui concerne l'ordre des mots et nous vérifions s'il s'agit réellement de la construction impersonnelle. Comme résultat, il nous reste 253 résultats qui correspondent à la construction recherchée.

4. Les résultats

Les résultats sont affichés sous forme de concordancier qui présente les résultats à l'aide d'un tableau avec les colonnes indiquant le numéro du résultat, le contexte gauche, l'occurrence distinguée en rouge et en gris et le contexte droit :

576 occurrences.		Page: 1 >	
▼ N°	▼ Contexte gauche: 100 mots	▼ Occurrence:	▼ Contexte droit: 100 mots
<input type="checkbox"/> 1	Mais nous montrerons ici	que s' il est vrai	qu' elles représentent un vivier foncier dans certaines localités , elles ne sont pas toujours faciles à prendre ni à mettre en valeur .
<input type="checkbox"/> 2		Il n' est pas possible de parler de quelque objet que	ce soit sans le qualifier , ne serait -ce que pour le différencier ou pour en souligner un caractère , au fil du propos .
<input checked="" type="checkbox"/> 3	Récusation classique au XIX siècle chez les critiques de l' enseignement , à laquelle correspond la revendication progressiste pour une éducation intégrale .	Il n' est pas étonnant que	les principaux propagandistes de cette éducation aient souvent affirmé qu' elle était la seule rationnelle
<input checked="" type="checkbox"/> 4		Il est possible aussi que	l' étudiant possède l' argent nécessaire grâce à un job qui cependant peut le priver d' une autre ressource nécessaire à la mobilité : le temps .
<input type="checkbox"/> 5	Ces étudiants ont conscience	qu' il est déjà très difficile	pour leurs parents de fournir chaque mois l' argent nécessaire , ce qui les dissuade fortement de dépenser pour faire des visites , pour être mobile : " quand t' as pas assez d' argent , tu vas te dire bon , je vais d' abord payer mon loyer puis tu vas pas te dire ouais je vais mettre 120 euros pour un aller retour , je ne sais pas pour aller visiter le sud de la France quoi , tu vois , c' est un peu ça quoi donc l' argent ... quand t' as l' argent , c' est vrai
<input checked="" type="checkbox"/> 6		Il n' est pas sûr pourtant qu'	il faille opérer une différenciation aussi nette que le suggère Balacheff entre le caractère global de la coutume qui régulerait le fonctionnement social de la classe dans la durée , et le caractère local du contrat didactique , compris comme un système d' obligations réciproques spécifique du contenu de la connaissance visée (Brousseau , 1998) .
<input type="checkbox"/> 7		Il serait pourtant erroné de croire que	cet ordre nouveau relève de la seule responsabilité et de la seule autorité du maître .
<input checked="" type="checkbox"/> 8		Il est vraisemblable que	les individus qui entrent tardivement dans le panel (à partir de 1978 , c' est-à-dire 32 ans pour la génération 1946) cotisaient auparavant auparavant à un autre régime .
<input type="checkbox"/> 9	10 À ceci près	qu' il est difficile	de comparer le bien -être procuré par l' inactivité au bien -être procuré par une retraite " normale " .
<input type="checkbox"/> 10	Une des raisons pour lesquelles	il est difficile d' identifier avec précision l' impact salarial de la formation continue en France serait qu'	on observe chaque année un mélange de véritables et de fausses actions de formation .

Figure 4 : Les résultats Scientext

Nous pouvons décocher les résultats qui ne sont pas pertinents pour notre recherche. Ils deviennent alors gris. Nous pouvons changer également la manière

d'afficher les résultats en choisissant le nombre de mots du contexte : il peut aller de 5 mots jusqu'à 100 mots. Si nous avons besoin d'informations supplémentaires sur un résultat précis, nous pouvons y accéder en cliquant sur l'occurrence. Nous voyons alors le contexte plus large et, si nous le souhaitons, nous pouvons voir l'analyse syntaxique de toute la phrase où se trouve notre construction. L'analyse syntaxique nous fournit les informations concernant les dépendances syntaxiques entre les éléments particuliers, les lemmes et les formes apparaissant dans la phrase, la catégorie à laquelle appartiennent un mot et sa flexion. Une fois la recherche finie, nous pouvons la sauvegarder sous le format CSV avec le codage UTF-8.

Outre les résultats sous forme de concordancier, Scientext nous donne accès à des statistiques. Nous pouvons avoir la liste des lemmes qui sont classés sur la base de leur fréquence dans le corpus. Nous pouvons également étudier la répartition des lemmes en ce qui concerne la partie du texte (introduction, développement, conclusion, annexe, notes, résumé) et la discipline scientifique ; Scientext nous présente leur nombre absolu et le nombre relatif d'occurrences.

5. Le corpus de contraste

Comme corpus de contraste, nous avons choisi un autre corpus associé au site Scienquest, à savoir l'Est Républicain qui contient 5 655 742 mots (22 textes). Ce corpus contient des textes puisés dans un corpus plus grand, issu de la collaboration entre le journal *L'Est Républicain* et le CNRTL qui est composé des fichiers, chacun comprenant l'ensemble d'éditions du journal. *L'Est Républicain* est un journal local composé de trois parties principales: pages internationales et nationales communes à toutes les éditions locales, pages régionales communes aux éditions locales d'une région (Vosges, Lorraine), pages locales qui sont spécifiques de l'édition locale (Nancy, Pont-à-Mousson etc.). Kamel Nebhi et Bertrand Gaiffe ont encodé le corpus en y ajoutant les modules *core*, *textstructure*, *header* et *figure*. Achille Falaise, de l'équipe Scientext, après avoir choisi 22 textes de ce corpus, les a annotés avec les logiciels MElt et Malt. Grâce à cela, le corpus est analysé morphosyntaxiquement et au plan syntaxique. En ce qui concerne notre recherche, nous avons particulièrement apprécié le fait que les corpus au sein du Scientext soient analysés syntaxiquement; ce qui nous a permis de vérifier nos hypothèses. De fait, le corpus annoté morphosyntaxiquement et syntaxiquement donne au chercheur de nouvelles possibilités d'exploitation : il permet de rechercher des constituants de la phrase par parties

de discours et par lemmes en établissant les relations syntaxiques entre des éléments; ils permettaient également de préciser les traits morphosyntaxiques. L'annotation morphosyntaxique et syntaxique élargit donc les possibilités de la linguistique de corpus en permettant d'axer la recherche sur l'aspect grammatical, ce qui est essentiel dans notre recherche.

Le corpus l'Est Républicain est donc un corpus sur qui nous répéterons les mêmes requêtes que celles que nous avons effectuées grâce au corpus Scientext. Ainsi, nous connaissons la fréquence des mêmes constructions dans le corpus du langage scientifique et le corpus de contraste. Cela nous permettra de tirer des conclusions concernant la particularité des constructions dans l'écrit scientifique.

Le mode de recherche sur l'Est Républicain ressemble à celui de Scientext. Nous commençons notre recherche avec le mode de recherche libre en tapant les formes ou les lemmes ou les catégories des parties successives et en établissant les relations syntaxiques:

Relations syntaxiques			
Mot 1	sujet de (suj)	Mot 2	X
Mot 3	objet direct de (obj)	Mot 2	X

Figure 5 : La recherche libre sur le corpus l'Est Républicain.

Pour que les résultats soient fiables, nous avons essayé d'être le plus proches possible des requêtes effectuées sur Scientext. Dans ce but, nous passons à la recherche avancée et nous affinons notre requête. Nous créons une variable contenant les mêmes listes des lemmes extraits grâce à Scientext et nous l'insérons dans notre requête, p.ex. :

\$ver = échapper, agir, aller, apparaître, appartenir, arriver, avérer, convenir, découler, demeurer, devenir, devoir, donner, ensuivre, exister, faire défaut, faire, falloir, importer, manquer, paraître, plaire, poser, pouvoir, provenir, réaliser, résulter, ramener, reconnaître, recourir, ressortir, rester, sembler, sortir, suffire, valoir

Main = <form=il|Il,,#0>&&<lemma=\$ver,,#1>&&<lemma=que,,#2> :: (suj,#1,#0)
(obj,#1,#2)

Pourtant, étant donné que chaque corpus était analysé syntaxiquement par un outil différent, nous avons été contraints d'ajuster notre requête aux caractéristiques d'un corpus donné. À cause de cela, nous pouvons rencontrer des différences au niveau des relations syntaxiques.

Les résultats du l'Est Républicain sont affichés sous forme de concordancier, ce qui nous a aidés dans la fouille manuelle. De fait, le contexte nous a souvent permis de décider si la construction trouvée convergeait avec la construction recherchée.

Une fois les résultats des deux corpus obtenus, nous les avons comparés dans un tableau. Nous avons confronté le nombre d'occurrences, ainsi que la fréquence relative (le nombre d'occurrences divisé par le nombre des mots), qui correspond au nombre de résultats pour la proportion sur 10 000 mots. Les résultats quantitatifs, ainsi que les résultats qualitatifs sont présentés dans les chapitres suivants.

Partie 3

-

Les résultats et l'observation des constructions

Chapitre 4. Les résultats quantitatifs

Dans ce chapitre, nous allons présenter les résultats obtenus grâce à notre recherche sur les corpus. Le tableau suivant présente les résultats : nous y voyons le nombre d'occurrences et la fréquence relative d'une construction dans le corpus scientifique Scientext et dans le corpus de contraste l'Est Républicain.

			Corpus scientifique		Corpus de contraste		Ratio
Construction		Exemple	Le nombre d'occurrences	La fréquence relative (le nombre d'occurrences/le nombre des mots)	Le nombre d'occurrences	La fréquence relative (le nombre d'occurrences/le nombre des mots)	Corpus scientifique/corpus de contraste
Construction impersonnelle adjectivale	Il + V + ADJ + que	<i>Il est nécessaire que</i>	267	0,82	163	0,29	2,83
	Il + V + ADJ + de	<i>Il est indispensable de</i>	1069	3,27	371	0,66	4,95
		Total	1336	4,09	534	0,94	4,35
Construction impersonnelle verbale	Il + V + que	<i>Il semble que</i>	1056	3,23	305	0,54	5,98
	Il + V + de	<i>Il convient de</i>	1480	4,53	457	0,81	5,59
		Total	2536	7,77	762	1,35	5,76
Total impersonnel			3872	11,86	1296	2,29	5,18
	Passif réduit	Résultats obtenus	19431	59,51	14801	26,17	2,27
Constructions	Pronominal	<i>La longueur se</i>	2150	6,58	1215	2,15	3,06

pronominales passives	passif simple	<i>calcule...</i>					
	Pronomnal passif composé	<i>Ce lien s'est fait à travers des théories sur...</i>	240	0,74	358	0,63	1,17
		Total	2390	7,32	1573	2,78	2,63
Constructions passives impersonnelles	Passif impersonnel présent	<i>Il est admis que...</i>	161	0,49	46	0,08	6,13
	Passif impersonnel passé	<i>Il a été démontré que...</i>	41	0,13	19	0,03	4,33
		Total	202	0,62	65	0,11	5,64
Autres constructions	Infinitif passif	<i>Etre considéré</i>	3192	9,78	1478	2,61	3,75
	Pour + inf.	<i>Pour obtenir</i>	5462	16,73	11220	19,84	0,84
	Gérondif	<i>En utilisant</i>	3740	11,45	4401	7,78	1,47
	Impératif en tête de la phrase	<i>Notons</i>	626	1,92	388	0,69	2,78
	Prép. + lequel	<i>Selon lequel</i>	2140	6,55	1151	2,04	3,21
	Verbes modaux + inf.		10500	32,16	12691	22,44	2,83

Tableau 1 : L'occurrence réelle et relative des constructions

Nous observons que les constructions les plus fréquentes à l'écrit scientifique, selon leur fréquence relative, sont : le passif réduit, verbes modaux suivis de l'infinitif, *pour* suivi de l'infinitif, l'impersonnel et le gérondif. Pourtant, pour avoir une vision plus claire des caractéristiques de l'écrit scientifique, nous avons observé les constructions les plus spécifiques de l'écrit scientifique, en utilisant le ratio entre le langage scientifique et langage journalistique. Dans ce but, nous avons divisé la fréquence relative de chaque construction du corpus Scientext par la fréquence relative de la même construction dans le corpus des écrits scientifiques du Scientext.

La comparaison du ratio entre les deux corpus démontre que les constructions les plus spécifiques de l'écrit scientifique sont les constructions impersonnelles, les constructions passives impersonnelles et les constructions infinitives passives (les trois constructions prises séparément ce sont : le passif impersonnel présent, *il + V + que*, *il + V + ADJ + de*). Ce qui est curieux toutefois est que la construction qui nous avait semblé très typique pour l'écrit scientifique, à savoir, *pour* suivi de l'infinitif est la moins typique pour l'écrit scientifique des constructions étudiées. Dans ce qui suit, nous avons choisi de nous concentrer sur quelques constructions impersonnelles et passives infinitives.

Chapitre 5. La description des constructions

Dans ce chapitre, nous présenterons la description de quelques constructions exemplaires selon le modèle présenté dans la partie méthodologique. Nous nous concentrerons sur les constructions les plus représentatives selon l'analyse quantitative, à savoir les constructions impersonnelles et passives.

1. *L'impersonnel*

Comme caractéristique principale des constructions à l'impersonnel, nous pouvons indiquer l'absence de la catégorie de la personne qui diminue l'agentivité dans le procès exprimé. Ainsi, dans l'écrit scientifique, cela permet au chercheur de mettre accent sur des actions scientifiques plutôt que sur les agents. De plus, dans certaines situations, ce type de constructions permet de cacher son propre avis derrière la construction impersonnelle. Autrement dit, elle permet d'exprimer une information subjective d'une manière qui semble objective. La deuxième fonction des constructions impersonnelles est de nature discursive. De fait, elles permettent de changer le sujet en introduisant un nouveau ou en attirant l'attention du lecteur à de nouveaux aspects. Dans notre analyse des constructions, nous avons pris en compte quatre constructions impersonnelles : *il V ADJ que*, *il V ADJ de*, *il V que*, *il V de*.

A. **Il + V + ADJ + que**

C'est une construction impersonnelle à sujet phrastique extraposé, composée du pronom impersonnel *il* (appelé également le sujet pléonastique ou explétif) suivi du verbe copule (*être*, *sembler*, *rester*, *apparaître*, *devenir*, *demeurer*, *paraître*), un adjectif et la conjonction *que* qui introduit le complément phrastique. Les adjectifs qui font partie de cette construction dans les corpus scientifiques sont: *absurde*, *acceptable*, *banal*, *bon*, *clair*, *commode*, *commun*, *difficile*, *erroné*, *étonnant*, *évident*, *excessif*, *fréquent*, *fructueux*, *important*, *impossible*, *indéniable*, *indispensable*, *inintéressant*, *intéressant*, *inutile*, *légitime*, *mauvais*, *nécessaire*, *optimal*, *piquant*, *possible*, *probable*, *radical*, *rare*, *remarquable*, *sûr*, *significatif*, *souhaitable*, *théorique*, *utile*, *vrai*, *vraisemblable*.

- 1) La construction remplit une importante fonction rhétorique en se composant avec les adjectifs comme *vrai*, *clair*, *remarquable*, *certain*, *évident*, *indéniable*, *sûr*. De fait, elle sert alors à donner de la valeur axiomatique aux thèses de l'auteur de l'article (1). Dans

ce cas-là, le verbe copule est à l'indicatif présent et le complément phrastique est à l'indicatif :

- (1) **Il est vrai que** la rareté des exemples et les mises en garde de l'auteur sur le caractère non **confessionnel** de la majorité de ces conflits empêchent d'énoncer des conclusions hâtives. [his-art-135-body]

Elle permet également de présenter les résultats de la recherche. Il n'est donc pas étonnant que sa fréquence relative soit la plus élevée dans la conclusion :

- (2) **Il est également évident que** la nature de la langue dans les interactions médiatisées, un peu comme l'oral à l'écrit, est intimement liée à la dimension socio-cognitive [psy-hdr-647-conclusion] (2)

- 2) Dans le cas des adjectifs à valeur épistémique : *envisageable, probable, concevable, possible, impossible, vraisemblable*, la construction sert principalement à introduire l'hypothèse. Étant donné qu'elles expriment la possibilité, les têtes sémantiques, dans ce cas, exigent que le complément phrastique soit au subjonctif, le mode présentant le procès comme une conception de l'esprit, pas située sur le plan de la réalité. (Martinet, 1979, p.120) La construction est très souvent accompagnée par un adverbe qui diminue ou renforce la probabilité tel que *pratiquement, fort, très, infiniment, parfaitement, hautement, peu, plus, moins, difficilement, assez, tout à fait*, et/ou un adverbe, une conjonction qui introduit ou justifie l'introduction de l'hypothèse : *donc, d'ailleurs, également, toutefois, en effet, même, cependant, ainsi, pourtant, aussi* (3) (4). Dans ce cas, le complément phrastique est au subjonctif.

- (3) **Il est fort probable toutefois qu'**une grande partie de celles-ci soit des femmes ayant définitivement cessé toute activité, et pour qui le problème de l'âge minimal pour avoir 40 annuités de cotisation ne se pose pas réellement. [eco-art-141-body]. (3)

- (4) **Il est donc probable que** ces contrats agricoles jouent un rôle non négligeable dans l'allocation du risque. [eco-art-141-body] (4)

- 3) La construction donne l'ordre, exprime une nécessité si elle est accompagnée par les adjectifs (modalité déontique) : *capital, nécessaire, important, souhaitable, utile, inutile, indispensable*. Le verbe copule est le plus souvent à l'indicatif présent. Néanmoins, il apparaît également au conditionnel présent et à l'imparfait. Le complément phrastique est alors au subjonctif. Sa fréquence relative est plus élevée dans l'introduction où elle justifie la légitimité de la recherche (5).

(5) Ensuite, le phénomène de l'ambiguïté lexicale est caractéristique des langues et, si l'on s'inscrit dans une perspective où les applications en TAL doivent demander une adaptation minimale de la part des utilisateurs, alors **il apparaît nécessaire que** ces applications soient capables de traiter l'ambiguïté lexicale, ce qui suppose une modélisation préalable. [tal-art-626-introduction] (5)

4) Enfin, la construction *il V ADJ que* sert à l'évaluation : elle permet de juger les phénomènes, les événements en présentant son jugement sous une forme de construction impersonnelle. Elle permet également de souligner l'importance d'un événement, d'attirer l'attention sur un événement. En ce qui concerne la fonction discursive, elle permet d'introduire un nouveau thème. Elle est alors accompagnée par les adjectifs : *étonnant, rare, normal, absurde, acceptable, banal, bon, commode, commun, difficile, erroné, excessif, fréquent, fructueux, inintéressant, intéressant, légitime, mauvais, optimal, piquant, radical, significatif, théorique* et le complément phrastique est au subjonctif (6) (7):

(6) **Il est significatif que** pour photographier la société française du début des années quatre-vingt, un organisme aussi officiel que la Datar (Délégation à l'aménagement du territoire et à l'activité régionale) fasse appel à des photographes d'art et non à des documentaristes [Paysages, photographies. La Mission photographique de la Datar, 1985]. [ant-art-265-body] (6)

(7) **Il est étonnant que** biodiversité n'apparaisse pas dans les dictionnaires spécialisés puisqu'il s'agit apparemment d'une discipline à part entière au sein des sciences de la nature et de l'environnement. [lin-the-657-body] (7)

B. Il + V + ADJ + de

C'est une construction impersonnelle à sujet phrastique extraposé, composée du pronom impersonnel *il* suivi d'un verbe copule, un adjectif et la préposition *de* qui introduit le verbe à l'infinitif. La diversité des adjectifs est beaucoup plus grande que dans la construction *il + V + ADJ + que*, mais elle ne prend que cinq verbes copules : *être, sembler, devenir, paraître, rester*. Plusieurs valeurs sémantiques peuvent être dégagées.

1) La construction indique la probabilité de réalisation. Elle sert principalement à introduire l'hypothèse et exprimer les restrictions et les possibilités de la recherche. Comme la construction en *que*, la construction est très souvent accompagnée par un adverbe ou une locution adverbiale qui diminue ou renforce la probabilité tel que *pratiquement, sans doute, peut-être, peu, totalement, tout à fait, certes*. Par ailleurs, on observe également des

adverbes qui restreignent la possibilité en ce qui concerne son aspect temporel : *actuellement, souvent, parfois, rarement, toujours, pas encore* et/ou un adverbe, une conjonction qui introduit ou justifie l'introduction de l'hypothèse : *en effet, également, donc, en revanche, alors, tout de même, cependant, par conséquent, néanmoins, d'ailleurs, encore, ainsi* (8). Elle contient très souvent le pronom personnel *nous* ou, plus rarement, *leur, lui*. (9).

(8) **Il est donc possible de** négliger toutes ces variations, non seulement en raison de leur faible ampleur, mais encore parce qu'elles se neutralisent ; aussi sera-t-il possible de travailler, dans l'approche du phénomène du jugement de paraphrase dans le corpus, sur des nombres plus grands que si l'on prenait toutes les catégories séparément. [sed-the-317-body] (8)

(9) De fait, **il nous est impossible de** tester la médiation de l'effet du conflit et des buts sur l'apprentissage, par le mode de régulation du conflit. [psy-the-504-body] (9)

2) La construction *il V ADJ de* exprime une nécessité. Elle permet très souvent au chercheur d'introduire un thème d'analyse, d'expliquer son procès de raisonnement, de justifier ses choix lors d'analyse. C'est pourquoi la construction peut être trouvée dans tout l'article sans préférence pour une partie particulière. La construction est alors très souvent accompagnée par un adverbe ou une locution adverbiale qui justifie, explique la nécessité : *aussi, en effet, donc, toutefois, également, alors, en outre, ensuite, pourtant, par conséquent* ou souligne son importance : *particulièrement, très, tout à fait*.

L'infinitif qui suit la préposition *de* appartient très souvent aux verbes de la construction du modèle d'analyse ou d'analyse elle-même. Dans le premier cas, il démontre alors un thème d'analyse : *se centrer sur, prendre en compte, interroger* (10). Dans le deuxième cas, l'infinitif concerne la description des données ou l'interprétation des résultats, ce sont alors des verbes d'examen, de constat, d'évaluation tels que : *approfondir, compléter, définir, expliciter, effectuer, observer, analyser, expliquer, rappeler* (11). Le verbe copule utilisé dans cette construction est uniquement *être*. Le plus souvent, il est à l'indicatif présent ; pourtant, nous le pouvons aussi rencontrer au conditionnel présent, au futur simple ou à l'imparfait.

(10) **Il est donc nécessaire d'envisager** plusieurs niveaux d'organisation textuelle et leurs modes d'articulation. [lin-hdr-364-body] (10)

- (11) C'est pourquoi **il est crucial de définir** les conditions de possibilité d'une enquête fondée sur le rapport spéculaire entre un groupe professionnel et sa visibilité médiatique. [infcom-art-5-body] (11)

3) La construction *il V ADJ de*, tout comme la construction en *que*, permet de juger les phénomènes, les événements. Pourtant, elle admet un plus grand nombre d'adjectifs tels que : *difficile, facile, utile, aisé, préférable, rare, pertinent, délicat, vain, légitime, inutile, optimal, illusoire, loisible, économique, naturel, erroné, étonnant, révélateur, absurde, profitable, logique, fréquent, hâtif, hasardeux, juste, bon, faux, absent, normal, réducteur, idiot, symptomatique, instructif, commode, immédiat, injuste, paradoxal, usuel, indifférent, correct, cohérent, illégal, anodin, utopique, prématuré, banal, inintéressant, radical, piquant, remarquable, impensable, aléatoire, inexact, illogique, exact*. Ces adjectifs peuvent être divisés en adjectifs axiologiques (*pertinent, logique, utile*), portant sur la temporalité (*rare, fréquent*), sur la complexité (*difficile, facile*). Pourtant, la plupart d'entre eux ce sont des adjectifs évaluatifs qui n'entrent pas dans des catégories plus spécifiques : *correct, utopique, banal, loisible*. Outre la fonction d'évaluation, elle remplit une fonction discursive en permettant d'introduire un nouveau thème (13). La construction est parfois accompagnée par le pronom personnel *nous*. Le verbe copule (*être, sembler, devenir, rester, paraître, apparaître*) est le plus souvent au présent de l'indicatif, au conditionnel présent ou à l'imparfait.

- (12) **Il est difficile de dire** combien de textes exactement sont concernés ; il faudrait pour cela fixer une limite, arbitraire, au-delà de laquelle décider que l'usage des marques personnelles et du *il* impersonnel ou du passif est à considérer comme particulièrement élevé. [lin-the-115-body] (12)

- (13) En premier lieu, **il est utile de faire le point** sur les défauts structurels de la gestion publique française. [his-art-166-body] (13)

C. **Il + V + que**

C'est une construction impersonnelle qui se compose du pronom impersonnel *il* (le sujet apparent) suivie d'un verbe à la troisième personne du singulier et la conjonction de subordination *que* qui introduit la subordonnée complétive. Plusieurs types d'emploi se dégagent.

- 1) La construction sert à informer de l'existence d'un fait, avec plus (*rester, arriver, s'avérer*) ou moins (*sembler, se pouvoir, apparaître, paraître*) de certitude (14). La

construction remplit aussi une fonction discursive : elle permet d'introduire un nouveau thème (15). Vu ces emplois, il n'est pas étonnant que la construction soit fréquente dans le développement et dans la conclusion de l'article. Le verbe est le plus souvent à l'indicatif présent, au conditionnel présent ou à l'imparfait. Le complément phrastique est soit à l'indicatif (*rester, demeurer*), soit au subjonctif (*arriver, sembler, se pouvoir, paraître*).

(14) Plus généralement, **il semblerait que** le critère d'identité de fonctionnement structural soit efficace pour circonscrire celles des variables phonologiques que nous avons appelées symétriques. [lin-hdr-396-body] (14)

(15) **Il apparaît que** les élèves non suivis se placent massivement dans les déciles 1 et 2, ce qui laisse présager un échec majeur de l'apprentissage de la lecture au CP. [sed-com-264-body] (15)

- 2) Si elle est accompagnée des verbes : *s'ensuivre, ressortir, résulter, découler*, la construction permet de tirer des conclusions de ce qui a été dit, d'un fait présenté. C'est pourquoi la construction est fréquente dans les conclusions. Une cause logique de la conclusion peut être exprimée au niveau de la construction même grâce au syntagme nominal introduit à l'aide de la préposition *de* (16). Dans le cas contraire, il arrive souvent que la construction contienne le pronom anaphorique *en* (17). Le complément phrastique est à l'indicatif.

(16) **Il ressort de ces travaux que** les comportements d'adhésion apparaissent comme fortement contraints par le contexte interactionnel dans la mesure où ces comportements ne sont pas introduits par l'enfant à n'importe quel moment de l'interaction, mais consécutivement à certaines sollicitations du parent [psy-art-382-body] (16)

(17) **Il en découle que** le programme d'optimisation de l'entreprise peut s'écrire. [eci-art-221-body] (17)

- 3) La construction, quand elle contient les verbes : *importer, falloir, convenir*, permet d'exprimer les nécessités méthodologiques de la recherche ou les conditions nécessaires pour qu'une hypothèse soit vraie. La construction est absente dans l'introduction. La construction admet un adverbe ou une locution adverbiale : *néanmoins, aussi, à tout le moins, donc, seulement, d'abord, également* qui marquent un rapport (de conséquence, de restriction) avec la préposition précédente. La construction peut contenir aussi l'objectif visé de l'introduction introduit par la préposition *pour*. Le complément phrastique est au subjonctif.

- (18) **Il importe seulement que** les données pragmatiques nécessaires à l'interprétation des énoncés ne court-circuitent pas les éléments que peut fournir l'analyse linguistique esquissée ci-dessus à propos de d'ailleurs. [lin-art-176-body] (18)
- (19) Cependant, **il faut également que** les traits distinctifs de la cible soient suffisamment saillants pour pouvoir être extraits de la stimulation. [psy-art-442-body] (19)

D. Il + V + de

C'est une construction impersonnelle qui se compose du pronom impersonnel *il* (le sujet apparent) suivie d'un verbe à la troisième personne du singulier et la préposition *de* qui introduit une proposition infinitive. Les verbes qui font partie de cette construction ne sont pas nombreux : *importer, suffire, convenir, s'agir*. Comme la construction en *que*, la construction a la valeur injonctive, mais celle-ci est moins générale. Elle sert principalement à exprimer les nécessités méthodologiques ou des approfondissements de la recherche suggérés à des recherches ultérieures. Il n'est donc pas étonnant que la conclusion soit la partie de l'article où la construction est relativement plus fréquente. Le verbe est le plus souvent à l'indicatif présent, au conditionnel présent, au futur simple ou à l'imparfait. Les verbes à l'infinitif qui suivent la préposition *de* sont essentiellement des verbes d'analyse (*montrer, établir, examiner, analyser, déterminer, construire, mesurer*). La construction peut contenir des adverbes qui soulignent la nécessité ou la présente comme une suite logique de ce qui a été dit comme *alors, certainement, d'abord, avant tout, donc, enfin, premièrement, simplement*.

- (20) Mais avant d'étudier les dynamiques actuelles, **il convient de retracer** les différentes étapes de construction urbaine afin de comprendre les spécificités des processus en cours. [geo-art-7-body] (20)
- (21) Si on accorde aux discours à visée scientifique (Granger, 1993) sur le social le même statut qu'aux discours sur la nature, le travail transpositif s'impose pour les enseignements économiques et sociaux, au même titre que pour les disciplines scientifiques. **Il importe alors d'en étudier** les spécificités. [sed-art-49-body] (21)

2. *Le passif*

Le passif traditionnel est une construction composée du nom, du verbe *être*, du participe passé d'un verbe transitif et, éventuellement, du complément d'agent.

Nous observons une abondance de formes passives dans l'écrit scientifique. Pour cela, dans cette partie de notre travail, nous allons réfléchir aux fonctions que peut remplir cette construction au sein de l'écrit scientifique. De fait, la plupart des linguistes s'écartent de plus en plus de la thèse de la grammaire transformationnelle selon laquelle le passif n'est qu'un équivalent de la proposition homologue à l'actif. Aujourd'hui, on constate que la phrase à la voix active et son homologue passif n'ont pas d'équivalence sémantique exacte et on trouve que le passif peut être utile :

lorsqu'on veut changer le profil communicatif d'un texte (le passif peut, par exemple, assurer une progression à thème constant; le complément devient, par sa position en tête de phrase, apte à fonctionner comme thème [:440]). Le passif est également une ressource stylistique importante puisqu'en permettant souvent d'effacer le sujet, il contribue à alléger la structure phrastique; il rend en outre possible de passer d'une proposition relative à des structures plus brèves, comme la participiale et l'épithète détachée (Riegel et al, cité par Gaatone, p. 237)

Une de ces fonctions est le fait remarqué par Riegel (1994), à savoir la possibilité d'élimination du premier argument du verbe, ce qui est appelé des passifs non achevés ou « tronqués ». Cette fonction avait été déjà remarquée par Gaatone D. en ce qui concerne les textes scientifiques, selon qui ce type de textes abonde en passifs tronqués (1998, p.214). La fonction d'évitement du premier argument comme évitement de la mention de l'« instance responsable » (Riegel et al., cité par Gaatone, p. 237) permet d'omettre ce qui est très souvent inutile dans le texte scientifique puisque l'identité de cette instance est dans la plupart des cas connue ou superflue (pour donner un exemple, l'auteur décrit tout le temps sa recherche dans le développement de l'article).

Le passif est une forme qui contribue à l'impersonnalité du texte technoscientifique (Kocourek 1991, p.85). Cela est visible dans les cas où le passif pourrait être dans certains cas remplacé avec la construction impersonnelle en ayant recours au pronom *on*:

(22) Comme on l'a souligné plus haut, avant la fin du XVIIIe siècle, **la diversité du vivant est regardée**, plus ou moins explicitement, comme l'actualisation mondaine d'une loi d'essence combinatoire.

« la diversité du vivant est regardée » – « on regarde la diversité du vivant ». Comme nous le savons, ce pronom est largement utilisé dans les articles scientifiques et ainsi cette construction pourrait être ambiguë au niveau de l'inclusion ou de l'exclusion de

l'énonciateur. Même si c'est vrai que le contexte pourrait écarter ces doutes, nous devons aussi insister sur la fonction stylistique du passif tronqué. Grâce à l'enlèvement du premier argument, la phrase devient moins lourde.

Une autre caractéristique du texte scientifique est le fait qu'en apparence on a affaire à un passif tronqué, mais les « agents », c'est-à-dire, le plus souvent, les auteurs, les partisans d'une thèse, les adhérents d'un courant sont souvent mentionnés dans des parenthèses.

- (23) **L' implication organisationnelle, et en particulier l' implication affective, est généralement liée** à une bonne performance dans l'entreprise et à une plus grande motivation (Meyer, Paunonen, Gellatly, Goffin et Jackson, 1989).

Le choix de la structure active ou passive est très souvent lié à la cohésion textuelle. Il arrive souvent qu'une information connue, le thème, soit préservée dans le sujet. Le passif permet donc de ne pas changer le sujet d'une proposition à l'autre en renvoyant au sujet précédent, ce qui est appelé par M. Riegel « isotopie référentielle » (cité par Gaatone, 1998, p.224). Ainsi, à côté des formes passives, on retrouve très souvent des éléments anaphoriques :

- (24) La première analyse se centre prioritairement sur le sens manifeste du contenu des discours des enseignants alors que la deuxième analyse se centre davantage sur les cooccurrences des mots relevés au sein de ces discours. **Ces deux types d'analyses ont été utilisés** dans le but d'enrichir l'analyse et l'interprétation des données puis, d'autre part, dans le but d'augmenter l'objectivité du regard porté sur les données déclarées par les enseignants. [sed-the-356-body]

Par manque de temps, nous avons restreint notre étude à un sous-ensemble de constructions passives. Pour notre analyse nous avons choisi les constructions passives accompagnées par un verbe modal, un autre élément typique à l'écrit scientifique.

A. Pouvoir + être + pp

C'est une construction passive qui se compose du verbe *pouvoir*, suivi du verbe *être* à l'infinitif, l'auxiliaire d'un participe passé. La construction est souvent accompagnée par un verbe d'association ou un verbe de cause: *associer, assimiler, lier, ramener, réduire, généraliser, comparer, rapporter, relier, corrélér, conduire* et permet de mettre en rapport deux phénomènes, le premier étant très souvent la conséquence du deuxième. Cela advient, grâce au verbe *pouvoir*, sous forme d'une hypothèse. Le verbe *pouvoir* est le plus souvent à l'indicatif présent. Il apparaît aussi au conditionnel présent, au futur simple et à l'imparfait.

- (25) Une partie des écarts **peut être liée à** la plus grande proportion de femmes (de longévité supérieure) dans la catégorie des bas salaires. [eco-art-9-body]
- (26) Mais ce résultat **pourrait être** simplement **dû à** un manque de précision dans les estimations, par exemple, à cause d'erreurs de mesure sur l'offre de travail. [eco-art-132-notes]

La construction permet d'estimer, de juger un phénomène en le comparant à un autre. Elle sert très souvent à définir, à préciser un terme, à tirer des conclusions. Les participes passé proviennent alors des verbes de compréhension (*voir, lire, percevoir, comprendre, appréhender, concevoir*) d'interprétation (*interpréter, décrire, définir, analyser, entendre, désigner*) ou d'évaluation (*considérer, envisager, traiter, penser, représenter, saisir*).

Présente dans toutes les parties de l'article, elle apparaît le plus fréquemment dans le développement et dans la conclusion.

- (27) Réciproquement, tout procès de sémiotisation **pourra être abordé comme** un *cours d'action*, rassemblant plusieurs protagonistes (éventuellement " fictifs "), captés par un certain " jeu " ou " genre " sémiotique. [ant-art-121-body]
- (28) La paramétrisation de l'utilité marginale de la consommation **peut être considérée comme** une manière de capturer une certaine forme d'hétérogénéité observable du facteur d'escompte des ménages ou de leur degré de patience. [eco-art-132-body]

3. Conclusions

Nous avons présenté des constructions impersonnelles et des constructions passives comme des exemples des constructions les plus emblématiques du langage scientifique écrit. Nous avons illustré notre modèle en énumérant les constituants de la construction en indiquant les éléments grammaticaux et les éléments sémantiques (les adjectifs, les verbes) qui en font partie. Nous avons indiqué les temps et les modes que prennent les verbes de la construction. Outre cela, nous nous sommes concentrée sur les fonctions discursives et pragmatiques des constructions et nous avons pu dégager quelques fonctions particulièrement présentes dans ce genre.

Il est apparu qu'une des fonctions les plus exploitées est la fonction informant sur l'état des choses : elle informe de l'existence de certains phénomènes soit sous forme d'évidence : il + V + ADJ + que (*il est évident que*), soit sous forme de conclusion, de résultat de recherche : il + V + que (*il s'ensuit que*).

La deuxième fonction particulièrement exploitée dans le langage scientifique concerne la probabilité. Elle est exprimée grâce aux constructions : il + V + ADJ + que (*il est probable que*), il + V + ADJ + de (*il est envisageable de*), Il + V + que (*il semble que*), pouvoir être + pp. (*il peut être dû*). Les constructions adjectivales exprimant la probabilité sont souvent accompagnées par des adverbes qui la renforcent ou la diminuent la probabilité. Dans le même but, les constructions verbales changent le mode de l'indicatif en conditionnel (ce qui renforce l'aspect hypothétique de construction).

La fonction suivante fréquente dans le langage scientifique est la fonction déontique. Elle apparaît dans les constructions : il + V + ADJ + de (*il est crucial de*), il + V + que (*il importe que*) il + V + de (*il convient de*), devoir + être + pp (*l'aspect devrait être approfondi*). La fonction sert à exprimer les nécessités méthodologiques de l'étude ou introduit une condition nécessaire pour que l'hypothèse posée soit vraie.

Une autre fonction qui ressort de notre travail est l'évaluation : il + V + ADJ + que (*il est significatif que*), il + V + ADJ + de (*il est utile de*). Dans l'écrit scientifique, à côté du jugement à propos des faits présentés, elle sert pour l'auteur à attirer l'attention du lecteur sur certains aspects ou sert à introduire un nouveau sujet.

Conclusion

La recherche sur la langue de spécialité concerne le plus souvent son aspect lexicologique en se concentrant sur la terminologie. Or, il arrive souvent qu'un genre présente également des caractéristiques de type grammatical. Dans notre travail, nous voulions rapprocher l'aspect grammatical et sémantique en essayant de repérer les constructions particulièrement représentées dans l'écrit scientifique, en empruntant la définition de construction des grammaires de constructions, à savoir en la considérant comme une paire conventionnalisée forme-sens.

Si l'on reconsidère les problématiques de notre recherche, nous avons bien trouvé des constructions typiques à l'écrit scientifique. Il s'est avéré que l'écrit scientifique a fréquemment recours aux constructions passives, aux constructions impersonnelles, aux constructions contenant des verbes modaux. Dans la présente recherche, pour repérer des constructions, nous nous sommes basée sur les recherches antérieures, mais nous avons aussi repéré des constructions manuellement en parcourant des articles scientifiques. Pour avoir des résultats plus fiables, la recherche devrait être répétée en ayant recours aux outils qui trouveraient automatiquement les suites des éléments grammaticaux répétées fréquemment. De plus, le corpus de référence devrait être plus important : il devrait contenir des textes appartenant aux autres genres que l'écrit journalistique, par exemple des textes littéraires et des textes utilitaires.

Une partie de notre travail consistait également à élaborer un modèle d'analyse. Nous avons voulu relier les aspects syntaxiques/grammaticaux aux aspects sémantiques. Pour cela, nous nous sommes appuyée sur le lexique-grammaire et les grammaires de construction (la théorie de Fillmore). Cependant, dans notre modèle, nous avons ajouté les fonctions discursives et pragmatiques qui ne sont pas présentes dans les courants susmentionnés. Il nous semble pertinent de prendre en compte l'aspect pragmatique, car cela pourrait servir aux étudiants étrangers qui prennent connaissance du langage scientifique écrit. Ainsi, à côté de la description formelle de la construction, ils seront conscients de ses emplois réels.

Pour conclure, nous sommes très contente d'avoir pu travailler sur ce sujet. Cela nous a permis d'explorer les questions comme les rapports entre le genre et la grammaire, tout en nous approchant aux théories des grammaires de construction ou du lexique-grammaire. L'aspect particulièrement intéressant de cette étude était l'exploitation des

corpus. Ce travail nous a montré des possibilités et des limites des outils dont les corpus ont été analysés par des analyseurs différents. La méthodologie que nous avons adoptée nous semble pertinente : repérer les constructions fréquentes dans l'écrit scientifique, avoir recours au corpus de référence pour comparer des fréquences relatives, la description des constructions en prenant en compte l'aspect pragmatique.

Le travail sur le corpus et la rédaction du mémoire nous ont permis d'acquérir plus de la rigueur intellectuelle et nous ont appris de mieux organiser notre travail. Ce mémoire nous a aussi démontré la façon de travail en linguistique, car nos travaux antérieurs concernaient la littérature. Nous sommes également contente d'avoir eu la possibilité de travailler sur notre mémoire en France, cela a amélioré notre connaissance de la langue et nous a fait découvrir une autre culture, y compris une autre culture universitaire.

Bibliographie

Adam, J.-M. (2011). *La linguistique textuelle. Introduction à l'analyse textuelle des discours*. Paris : Armand Colin, coll. Cursus.

Baker Collin, F. (2009). La sémantique des cadres et le projet FRAMENET : une approche différente de la notion de « valence ». *Langages*, 176, 32-49.

Bakhtin, M. M. (1986). *Speech Genres and Other Late Essays*. Austin : University of Texas Press.

Bakhtin, M. M. (2004). Dialogic Origin and Dialogic Pedagogy of Grammar: Stylistics in Teaching Russian Language in Secondary School, *Journal of Russian and East European Psychology*, 42 p.12-49.

Beacco, J.-C. (2004). Trois perspectives linguistiques sur la notion de genre discursif. *Langages*, 153, 109-119.

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers* (Vol. 23). Amsterdam : John Benjamins Publishing.

Branca-Rosoff, S. (1999). Des innovations et des fonctionnements de langue rapportés à des genres. *Langage et société*, 87, 115-129.

Calberg Challot, M. (2008). Quand un vocabulaire de spécialité emprunte au langage courant : le nucléaire, étude de cas. *Terminology*, 14(2), 183-203.

Desagulier, G. (2007). Quelques remarques sur l'origine des Grammaires des Constructions et le statut des représentations en linguistique. *HAL. Archives ouvertes*. Répéré à <https://halshs.archives-ouvertes.fr/halshs-00202008/document>

Desmet, I. (2006). Variabilité et variation en terminologie et langues spécialisées : discours, textes et contextes. *Communication aux septièmes journées scientifiques du réseau Lexicologie, terminologie, traduction (LTT) de l'Agence universitaire de la francophonie (AUF), Mots, termes et contextes*, Actualité scientifique, Paris : Editions des archives contemporaines, pp. 235-247.

Fillmore, Ch. J. (1976). Frame semantics and the nature of language. *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*. 280, 20-32.

Fløttum, K., Dahl, T. & Kinn, T. (2006). *Academic Voices: across languages and disciplines*. University of Bergen / Norwegian School of Economics and Business Administration. Amsterdam : John Benjamins Publishing Company.

François, J. (2008). Les grammaires de construction un bâtiment ouvert aux quatre vents. *Cahier du Crisco n° 26*. Caen : Centre de Recherches Inter-lingues sur la Signification en Contexte.

- Freyermuth, S. (2011). Un genre peut en cacher un autre : une histoire de détournement, *Linx*, 64-65. Répéré à <http://linx.revues.org/1411#quotation>
- Gaatone, D. (1988). *Le passif en français*. Bruxelles : Editions Duculot.
- Gaatone, D. (2006). À la recherche des fondements de la grammaire : les fonctions grammaticales et Le Bon Usage. *Revue belge de philologie et d'histoire*, 84, 929-952.
- Gardent, C., Guillaume, B., Perrier, G. & Falk, I. (2005). Maurice Gross' Grammar Lexicon and Natural Language Processing. *Proceedings of the 2nd Language and Technology Conference*. Poznań, Pologne.
- Gadet, F. (1996). Niveau de langues et variation intrinsèque. *Palimpsestes*, 10, 17-40.
- Gaudin, F. (1995). Champs, clôtures et domaines : des langues de spécialités à la culture scientifique. *Translators' Journal*, 40 (2), 229-237.
- Gross, M. (1975). *Méthodes en syntaxe. Le regime des constructions complétives*. Paris : Hermann.
- Gross, M. (1978). Correspondance entre forme et sens à propos du subjonctif. *Langue française*, 39, 49-65.
- Gross, M. (1986). *Grammaire transformationnelle du français, 1 - Syntaxe du verbe*. Paris : Cantilène.
- Grossmann, F. & Tutin, A. (2010). Les marqueurs verbaux de constat: un lieu de dialogisme dans l'écrit scientifique. Montpellier : Colloque international Dialogisme : langue, discours.
- Hyland, K. (2009). Corpus informed discourse analysis: the case of academic engagement. *Academic Writing: at the Interface of Corpus and Discourse*. London: Continuum, 110-128.
- Ibrahim, A. H. (2010). Supports d'actualisation et dualité constitutive du prédicat. *Supports et prédicats non verbaux dans les langues du monde*. Paris: CRL, 36-73.
- Kocourek, R. (1982). *La langue française de la technique et de la science*. John Benjamins Publishing Company.
- Krazem, M. (2011). Les genres de discours vus par la grammaire de la langue. *LINX* 64-65.
- Laporte, E. (2005). Graphes paramétrés et lexique-grammaire. *Interface lexique-grammaire et lexiques syntaxiques et sémantiques* (journée ATALA du 12 mars 2015).
- Lerat, P. (1995). *Les langues spécialisées*. Paris : PUF.
- Lyngfelt, B., Borin, L., Forsberg, M., Prentice, J., Rydstedt, R., Sköldberg, E., & Tingsell, S. (2012). Adding a Constructio n to the Swedish resource network of Språkbanken. In *Proceedings of KONVENS* (pp. 452-461).

- Mainguneau, D. (1993). *Le contexte de l'œuvre littéraire : Enonciation, écrivain, société*. Paris : Dunond.
- Martinet, A. & Martinet J. (1979). *Grammaire fonctionnelle du français*. Paris : Crédif.
- Mathieu, Y. Y. (2003). La grammaire de construction. *Linx*, 48. DOI : 10.4000/linx.121.
- Pecman, M. (2004). *Phraséologie contrastive anglais-français : analyse et traitement en vue de l'aide à la rédaction scientifique*. Thèse de doctorat. Dir. Henri Zinglé. Université de Nice-Sophia Antipolis.
- Petitjean, A. (1989). Les typologies textuelles. *Pratiques*. 62, 86-125.
- Puckica, J. (2008). Sens et relations de sens dans les grammaires de construction. *Lexis, E-Journal In English Lexicology*, n°1. Répéré à http://lexis.univ-lyon3.fr/IMG/pdf/Lexis_1_Puckica.pdf
- Rastier, F. (2001). *Arts et sciences du texte*. Paris : PUF.
- Rinck, F. (2010). L'analyse linguistique des enjeux de connaissance dans le discours scientifique. *Revue d'anthropologie des connaissances*, 4(3), 427–450.
- Spillner, B. (1992). Textes médicaux français et allemands. Contribution à une comparaison interlinguale et interculturelle. *Langages*, 105, 42-65.
- Swales, J.-M. (1990). *Genre Analysis. English in academic and research settings*. Cambridge University Press : Cambridge.
- Tran, T. T. H. (2010). *Aides logicielles à la rédaction universitaire. L'exemple des collocations*. Mémoire Master 2. Sciences du langage. Université Stendhal-Grenoble3 : Grenoble.
- Tutin, A. & Grossmann, F. (2013). *L'écrit scientifique : du lexique au discours*. Rennes : PUR.
- Tutin, A. (2010) : *Dans cet article, nous souhaitons montrer que...* Lexique verbal et positionnement de l'auteur dans les articles en sciences humaines. Enonciation et rhétorique dans l'écrit scientifique, *Lidil*, 41, 15-40.
- Tutin, A. (2007). Autour du lexique et de la phraséologie des écrits scientifiques. *Revue française de linguistique appliquée*, 12, 5-14.
- Pic, Elsa & Grégory Furmaniak. 2010. « Grammaire et degré de spécialisation ». *ASp* 58, 39-55.
- Yan, R. (2012). *Observation du lexique verbal transdisciplinaire et modélisation des patrons dans l'écrit scientifique : construction d'un dictionnaire électronique d'apprentissage d'encodage*. Mémoire Master 2. Université Stendhal - Grenoble 3, Grenoble.

Mathieu, Y. Y. (2009). La Grammaire de Construction. *Linx*, 48. Répéré à <http://linx.revues.org/121#quotation>

Sigles et abréviations utilisés

ADJ	l'adjectif
ADV	adverbe
Inf.	l'infinitif
Loc. adv.	La Locution adverbiale
Pp.	le participe passé
Pron.	le pronom
SN	syntagme nominal
Subj.	le subjonctif
V	le verbe

Table des illustrations¹

Figure 1. Choix de corpus	41
Figure 2. Recherche libre Scientext	42
Figure 3. Recherche avancée Scientext	42
Figure 4. Les résultats Scientext.....	43
Figure 5. La recherche libre sur le corpus l'Est Républicain	45
Tableau 1. L'occurrence réelle et relative des constructions	49
Tableau 2. La différence de l'occurrence relative entre le corpus Scientext et le corpus de contraste	51

Table des annexes²

Annexe 1 Les requêtes du Scientext	72
Annexe 2 Les requêtes de l'Est Républicain.....	82

Annexe 1

Les requêtes du Scientext

II + V + ADJ + que

\$adj = étonnant, évident, absurde, acceptable, banal, bon, clair, commode, commun, difficile, erroné, excessif, fréquent, fructueux, important, impossible, indéniable, indispensable, inintéressant, intéressant, inutile, légitime, mauvais, nécessaire, optimal, piquant, possible, probable, radical, rare, remarquable, sûr, significatif, souhaitable, théorique, utile, vrai, vraisemblable

Main = <form=/^(il|II)\$/,#1> && <tags=/VCONJ.+/,cat=V,#2> && <lemma=\$adj,#3> && <lemma=que,#4> :: (SUJ,#1,#2)(ATTS,#3,#2)(OBJ,#4,#2);

II + V + ADJ + de

\$adje = économique, égal, éloquent, équivalent, étonnant, étrange, évident, absent, absurde, agréable, aisé, aléatoire, analysable, anodin, arbitraire, avantageux, banal, beau, bon, capital, caricatural, clair, cohérent, commode, commun, congruent, constant, consubstantiel, correct, crucial, curieux, délicat, différent, difficile, douteux, effectif, envisageable, erroné, essentiel, exact, facile, faisable, fallacieux, faux, fondamental, fréquent, hâtif, habituel, hasardeux, homogène, idiot, illégal, illogique, illusoire, immédiat, impensable, important, impossible, imprudent, inévitable, inconcevable, incontestable, indéniable, indifférent, indirect, indispensable, inexact, influent, inintéressant, injuste, instructif, inutile, invraisemblable, juste, légitime, logique, loisible, long, minimal, modeste, nécessaire, négligeable, naturel, normal, notable, notoire, observable, optimal, paradoxal, pertinent, piquant, plausible, positif, possible, préférable, prématuré, présent, primordial, probable, productif, profitable, réducteur, révélateur, radical, raisonnable, rare, rationnel, remarquable, sûr, sage, significatif, souhaitable, spécifique, superflu, surprenant, susceptible, symptomatique, tolérable, typique, urgent, usuel, utile, utopique, vain, valable, vital, vrai, vraisemblable

Main = <form=/^(il|II)\$/,#1> && <tags=/VCONJ.+/,cat=V,#2> && <lemma=\$adje,#3> && <lemma=de,#4> :: (SUJ,#1,#2)(ATTS,#3,#2)(PREP,#4,#3);

Il + V + que

\$ver = échapper, agir, aller, apparaître, appartenir, arriver, avérer, convenir, découler, demeurer, devenir, devoir, donner, s'ensuivre, exister, faire défaut, faire, falloir, importer, manquer, paraître, plaire, poser, pouvoir, provenir, réaliser, résulter, ramener, reconnaître, recourir, ressortir, rester, sembler, sortir, suffire, valoir

Main = <form=/^(il|Il)\$/,#1> && <lemma=\$ver,#2> && <lemma=que,#3> ::
(SUJ,#1,#2)(OBJ,#3,#2);

Il + V + de

\$ver = échapper, advenir, agir, aller, apparaître, arriver, avérer, avoir lieu, convenir, débiter, découler, déduire, dépendre, devenir, desservir, devoir, ensuivre, falloir, importer, inscrire, intervenir, manquer, paraître, résulter, rendre, ressortir, rester, sembler, suffire, valoir, venir

Main = <form=/^(il|Il)\$/,#1> && <lemma=\$ver,#2> && <lemma=de,#3> &&
<tags=VINF,cat=V,#4> :: (SUJ,#1,#2)(PREP,#3,#2)(NOMPREP,#4,#3);

Le passif réduit

\$ver = priver, lier, donner, proposer, utiliser, élever, obtenir, associer, observer, mener, présenter, définir, réaliser, fonder, construire, étudier, produire, considérer, prendre, ouvrir, écrire, consacrer, effectuer, développer, faire, marquer, établir, retenir, poser, engager, percevoir, offrir, organiser, exprimer, publier, absoudre, induire, rencontrer, évoquer, choisir, décrire, interroger, constituer, intégrer, adopter, déterminer, acquérir, contraindre, finir, estimer, identifier, baser, envisager, fournir, créer, adapter, élaborer, supposer, employer, attendre, reconnaître, mettre en oeuvre, avancer, recueillir, exercer, analyser, éloigner, former, apporter, fixer, opérer, adresser, représenter, répéter, atteindre, comparer, soumettre, appeler, mettre, recevoir, concevoir, mentionner, fermer, inscrire, placer, émettre, intituler, viser, sélectionner, introduire, rechercher, juger, garantir, actualiser, réserver, citer, jouer, tirer, protéger, prévoir, constater, afficher, différer, rédiger, attribuer, croiser, désigner, mettre en place, enregistrer, tester, formuler, rendre, assurer, contrôler, pratiquer, ajouter, relier, tenir, calculer, enseigner, nommer, assister, saisir, inspirer, comprendre, déclarer, demander, envoyer, trouver, traiter, dominer, soutenir, contraster,

isoler, modifier, évaluer, défendre, extraire, hériter, négocier, appliquer, accepter, avancer, élaborer, résoudre, donner, réaliser, produire, mettre, ressentir, ouvrir, répéter, vendre, commettre, investir, proposer, produire

Main = <cat=N,#1> && <lemma=\$ver,tags=/Ppa.+/,cat=V,#2> :: (ADJ,#2,#1);

Passif pronominal

Présent

\$ver = faire, traduire, expliquer, écrire, développer, jouer, définir, opérer, tenir, observer, proposer, établir, interpréter, dessiner, dire, vouloir, mettre, maintenir, devoir, résumer, concevoir, déduire, rencontrer, lire, élaborer, prendre, porter, obtenir, déterminer, confirmer, offrir, saisir, falloir, consacrer, reconnaître, partager, accomplir, prononcer, impliquer, estimer, penser, lancer, réécrire, recruter, pratiquer, tisser, revendiquer, percevoir, conclure, distribuer, solder, préciser, risquer, satisfaire, doter, liser, repérer, vendre, suffire, fabriquer, décider

Main = <lemma=se,#1> && <lemma=\$ver,tags=/VCONJ.+/,cat=V,#2> :: (REF,#1,#2);

Passé

\$ver = faire, traduire, expliquer, écrire, développer, jouer, définir, opérer, tenir, observer, proposer, établir, interpréter, dessiner, dire, vouloir, mettre, maintenir, devoir, résumer, concevoir, déduire, rencontrer, lire, élaborer, prendre, porter, obtenir, déterminer, confirmer, offrir, saisir, falloir, consacrer, reconnaître, partager, accomplir, prononcer, impliquer, estimer, penser, lancer, réécrire, recruter, pratiquer, tisser, revendiquer, percevoir, conclure, distribuer, solder, préciser, risquer, satisfaire, doter, liser, repérer, vendre, suffire, fabriquer, décider

Main = <lemma=se,#1> && <lemma=être,#2> && <lemma=\$ver,cat=V,#3> ::
(REF,#1,#2)(AUX,#3,#2);

Le passif impersonnel présent

Main = <form=/^(il|Il)\$/,#1> && <lemma=être,#2> && <tags=/Ppa.+/,cat=V,#3> &&
<lemma=que,#4> :: (SUJ,#1,#2)(AUX,#3,#2)(OBJ,#4,#3);

Le passif impersonnel passé

Main = <form=/^(il|Il)\$/,#1> && <lemma=avoir,#2> && <form=été,#3> &&
<tags=/Ppa.+/,cat=V,#4> && <lemma=que,#5> ::
(SUI,#1,#2)(AUX,#3,#2)(AUX,#4,#3)(OBJ,#5,#4);

L'infinitif passif

Main = <form=/^(Être|être)\$/,#1> && <tags=/Ppa.+/,cat=V,#2> :: (AUX,#2,#1);

Pour + inf.

Main = <lemma=pour,#1> && <tags=VINFINF,cats=V,#2> :: (NOMPREP,#2,#1);

Gérondif

Main = <lemma=en,#1> && <cat=V,form=/ant\$/,#2> :: (NOMPREP,#2,#1);

Impératif en tête

Main = ^<form=/ons\$/,cat=V,#1>;

Prép + lequel

Main = <cat=O,#1> && <lemma=lequel,#2> :: (NOMPREP,#2,#1);

Verbe modal + inf.

Main = <lemma=/^(vouloir|pouvoir|devoir)\$/,#1> && <tags=VINFINF,cats=V,#2> ::
(OBJ,#2,#1);

Annexe 2

Les requêtes de l'Est Républicain

Il est ADJ que

\$adj = étonnant, évident, absurde, acceptable, banal, bon, clair, commode, commun, difficile, erroné, excessif, fréquent, fructueux, important, impossible, indéniable, indispensable, inintéressant, intéressant, inutile, légitime, mauvais, nécessaire, optimal, piquant, possible, probable, radical, rare, remarquable, sûr, significatif, souhaitable, théorique, utile, vrai, vraisemblable

Main = <form=il|Il,,#0>&&<cat=V,,#1>&&<lemma=\$adj,,#2>&&<lemma=que,,#3> ::
(suj,#1,#0) (ats,#1,#2)

Il est ADJ de

\$adje = économique, égal, éloquent, équivalent, étonnant, étrange, évident, absent, absurde, agréable, aisé, aléatoire, analysable, anodin, arbitraire, avantageux, banal, beau, bon, capital, caricatural, clair, cohérent, commode, commun, congruent, constant, consubstantiel, correct, crucial, curieux, délicat, différent, difficile, douteux, effectif, envisageable, erroné, essentiel, exact, facile, faisable, fallacieux, faux, fondamental, fréquent, hâtif, habituel, hasardeux, homogène, idiot, illégal, illogique, illusoire, immédiat, impensable, important, impossible, imprudent, inévitable, inconcevable, incontestable, indéniable, indifférent, indirect, indispensable, inexact, influent, inintéressant, injuste, instructif, inutile, invraisemblable, juste, légitime, logique, loisible, long, minimal, modeste, nécessaire, négligeable, naturel, normal, notable, notoire, observable, optimal, paradoxal, pertinent, piquant, plausible, positif, possible, préférable, prématuré, présent, primordial, probable, productif, profitable, réducteur, révélateur, radical, raisonnable, rare, rationnel, remarquable, sûr, sage, significatif, souhaitable, spécifique, superflu, surprenant, susceptible, symptomatique, tolérable, typique, urgent, usuel, utile, utopique, vain, valable, vital, vrai, vraisemblable

Main = <form=il|Il,,#0>&&<cat=V,,#1>&&<lemma=\$adje,,#2>&&<lemma=de,,#3> ::
(suj,#1,#0) (ats,#1,#2)

Il + V + que

\$ver = échapper, agir, aller, apparaître, appartenir, arriver, avérer, convenir, découler, demeurer, devenir, devoir, donner, ensuivre, exister, faire défaut, faire, falloir, importer, manquer, paraître, plaire, poser, pouvoir, provenir, réaliser, résulter, ramener, reconnaître, recourir, ressortir, rester, sembler, sortir, suffire, valoir

Main = <form=il|Il,,#0>&&<lemma=\$ver,,#1>&&<lemma=que,,#2> :: (subj,#1,#0)
(obj,#1,#2)

Il + V + de

\$ver = échapper, advenir, agir, aller, apparaître, arriver, avérer, avoir lieu, convenir, débiter, découler, déduire, dépendre, devenir, desservir, devoir, ensuivre, falloir, importer, inscrire, intervenir, manquer, paraître, résulter, rendre, ressortir, rester, sembler, suffire, valoir, venir

Main = <form=il|Il,,#0>&&<lemma=\$ver,,#1>&&<lemma=de,,#2> :: (subj,#1,#0)
(de_obj,#1,#2)

Le passif réduit

\$ver = priver, lier, donner, proposer, utiliser, élever, obtenir, associer, observer, mener, présenter, définir, réaliser, fonder, construire, étudier, produire, considérer, prendre, ouvrir, écrire, consacrer, effectuer, développer, faire, marquer, établir, retenir, poser, engager, percevoir, offrir, organiser, exprimer, publier, absoudre, induire, rencontrer, évoquer, choisir, décrire, interroger, constituer, intégrer, adopter, déterminer, acquérir, contraindre, finir, estimer, identifier, baser, envisager, fournir, créer, adapter, élaborer, supposer, employer, attendre, reconnaître, mettre en oeuvre, avancer, recueillir, exercer, analyser, éloigner, former, apporter, fixer, opérer, adresser, représenter, répéter, atteindre, comparer, soumettre, appeler, mettre, recevoir, concevoir, mentionner, fermer, inscrire, placer, émettre, intituler, viser, sélectionner, introduire, rechercher, juger, garantir, actualiser, réserver, citer, jouer, tirer, protéger, prévoir, constater, afficher, différer, rédiger, attribuer, croiser, désigner, mettre en place, enregistrer, tester, formuler, rendre, assurer, contrôler, pratiquer, ajouter, relier, tenir, calculer, enseigner, nommer, assister, saisir, inspirer, comprendre, déclarer, demander, envoyer, trouver, traiter, dominer, soutenir, contraster,

isoler, modifier, évaluer, défendre, extraire, hériter, négocier, appliquer, accepter, avancer, élaborer, résoudre, donner, réaliser, produire, mettre, ressentir, ouvrir, répéter, vendre, commettre, investir, proposer, produire

Main = <cat=N, #0><lemma=\$ver, #1>

Le pronominal passif:

Présent:

\$verb = fait, font, traduit, traduisent, explique, expliquent, écrit, écrivent, développe, développent, joue, jouent, définit, définissent, opère, opèrent, tient, tiennent, observe, observent, propose, proposent, établit, établissent, interprète, interprètent, dessine, dessinent, dit, disent, veut, veulent, met, mettent, doit, doivent, résume, résument, conçoit, conçoivent, déduit, déduisent, rencontre, rencontrent, lit, lisent, élabore, élaborent, prend, prennent, porte, portent, obtient, obtiennent, détermine, déterminent, confirme, confirment, offre, offrent, saisit, saisissent, faut, consacre, consacrent, reconnaît, reconnaissent, partage, partagent, accomplit, accomplissent, prononce, prononcent, implique, impliquent, estime, estiment, pense, pensent, lance, lancent, réécrit, réécrivent, recrute, recrutent, pratique, pratiquent, tisse, tissent, revendique, revendiquent, perçoit, perçoivent, conclut, concluent, distribue, distribuent, solde, soldent, précise, précisent, risque, risquent, satisfait, satisfaisent, dote, dotent, lit, lisent, repère, repèrent, vend, vendent, suffit, suffisent, fabrique, fabriquent, décide, décident

Main = <lemma=se, #0><form=\$verbe, #1>

Passé:

\$ver = faire, traduire, expliquer, écrire, développer, jouer, définir, opérer, tenir, observer, proposer, établir, interpréter, dessiner, dire, vouloir, mettre, maintenir, devoir, résumer, concevoir, déduire, rencontrer, lire, élaborer, prendre, porter, obtenir, déterminer, confirmer, offrir, saisir, falloir, consacrer, reconnaître, partager, accomplir, prononcer, impliquer, estimer, penser, lancer, réécrire, recruter, pratiquer, tisser, revendiquer, percevoir, conclure, distribuer, solder, préciser, risquer, satisfaire, doter, liser, repérer, vendre, suffire, fabriquer, décider

Main = <lemma=se, #0>&&<lemma=être, #1>&&<lemma=\$ver, #2> :: (aux_tps, #2, #1)

Le passif impersonnel :

Présent:

Main = $\langle \text{form}=\text{il}|\text{Il},\#0 \rangle \&\& \langle \text{lemma}=\text{être},\#1 \rangle \&\& \langle \text{cat}=\text{V},\#2 \rangle \&\& \langle \text{lemma}=\text{que},\#3 \rangle ::$
 $(\text{subj},\#1,\#0) (\text{mod},\#1,\#2)$

Passé:

Main =
 $\langle \text{form}=\text{il}|\text{Il},\#0 \rangle \&\& \langle \text{lemma}=\text{avoir},\#1 \rangle \&\& \langle \text{form}=\text{été},\#2 \rangle \&\& \langle \text{cat}=\text{V},\#3 \rangle \&\& \langle \text{lemma}=\text{que},\#4 \rangle ::$
 $(\text{subj},\#2,\#0) (\text{aux_tps},\#2,\#1)$

Infinitif passé

Main = $\langle \text{form}=\text{Être}|\text{être},\#0 \rangle \&\& \langle \text{cat}=\text{V},\#1 \rangle ::$ $(\text{mod},\#0,\#1)$

Pour + inf.

Main = $\langle \text{form}=\text{pour},\#0 \rangle \&\& \langle \text{cat}=\text{V},\#1 \rangle ::$ $(\text{obj},\#0,\#1)$

Gérondif

Main = $\langle \text{form}=\text{en},\#0 \rangle \&\& \langle \text{cat}=\text{V},\#1 \rangle ::$ $(\text{obj},\#0,\#1)$

Impératif en tête

Main = $\langle \text{form}=[\text{A-Z}]^+.\text{ons}/,\text{cat}=\text{V},\#1 \rangle$

Prép + lequel

Main = $\langle \text{cat}=\text{P},\#0 \rangle \&\& \langle \text{lemma}=\text{lequel},\#1 \rangle ::$ $(\text{obj},\#0,\#1)$

Verbe modal + inf.

\$mod = vouloir, pouvoir, devoir

Main = $\langle \text{lemma}=\text{\$mod},\#0 \rangle \&\& \langle \text{cat}=\text{V},\#1 \rangle ::$ $(\text{obj},\#0,\#1)$

Annexe 3

Les tableaux d'analyse

1. *Il V ADJ que*

La structure	Il + V + ADJ + que + ind./subj.
La définition	<p>Il [sujet apparent] + verbe copule + adjectif [l'attribut du sujet] + que + proposition [indicatif ou subjonctif]</p> <p>Une construction impersonnelle à sujet phrastique extraposé, composée du pronom impersonnel <i>il</i> (appelé également le sujet pléonastique ou explétif) suivi d'un verbe copule, un adjectif et la conjonction de subordination <i>que</i> qui introduit le complément phrastique.</p>
Les constituants et leurs propriétés morphologiques	<p>Il – pronom impersonnel, m, sg</p> <p>Verbe, 3 per, sg, indicatif</p> <p>Adjectif, m, sg</p> <p>Que – conjonction de subordination</p>
La tête lexicale	<p>ADJ : étonnant, évident, absurde, acceptable, banal, bon, clair, commode, commun, difficile, erroné, excessif, fréquent, fructueux, important, impossible, indéniable, indispensable, inintéressant, intéressant, inutile, légitime, mauvais, nécessaire, optimal, piquant, possible, probable, radical, rare, remarquable, sûr, significatif, souhaitable, théorique, utile, vrai, vraisemblable</p> <p>VERB : être, sembler, rester, apparaître, devenir, demeurer, paraître</p>
Les fonctions sémantiques et rhétoriques	
Modalité épistémique : évidence (vrai, clair, remarquable, certain, évident, indéniable,	

indispensable, sûr)	
Exemples	Commentaire
<p>Il est vrai que la rareté des exemples et les mises en garde de l'auteur sur le caractère non confessionnel de la majorité de ces conflits empêchent d'énoncer des conclusions hâtives. [his-art-135-body]</p> <p>Il est en effet clair que ce ne sont pas toutes les données qui vont être pertinentes à chaque fois. [lin-art-38-body]</p> <p>Cependant, il est évident que les électeurs d'origine étrangère votent davantage pour des politiciens d'origine étrangère que ne le font la majorité des électeurs. [spo-art-153-body]</p>	<p>La construction sert à donner une valeur axiomatique aux thèses de l'auteur de l'article ou elle permet d'introduire un nouveau sujet. Elle permet également de présenter les résultats de la recherche.</p> <p>La fréquence relative est la plus élevée dans la conclusion.</p> <p>Le verbe est à l'indicatif présent.</p> <p>Le complément phrastique est à l'indicatif.</p>
Modalité épistémique : probabilité (envisageable, probable, concevable, possible, impossible, vraisemblable)	
Exemples	Commentaire
<p>Il est fort probable toutefois qu'une grande partie de celles-ci soit des femmes ayant définitivement cessé toute activité, et pour qui le problème de l'âge minimal pour avoir 40 annuités de cotisation ne se pose pas réellement. [eco-art-141-body].</p> <p>Il est donc probable que ces contrats agricoles jouent un rôle non négligeable dans l'allocation du risque. [eco-art-141-body]</p> <p>Il est fort probable toutefois qu'une grande partie de celles-ci soit des femmes n'a pas validé 40 annuités avant l'âge de 60</p>	<p>La construction sert principalement à introduire l'hypothèse</p> <p>Le complément phrastique est au subjonctif.</p> <p>La construction est très souvent accompagnée par un adverbe qui diminue ou renforce la probabilité tel que <i>pratiquement, fort, très, infiniment, parfaitement, hautement, peu, plus, moins, difficilement, assez, tout à fait</i>, et/ou un adverbe, une conjonction qui introduit ou justifie l'introduction de l'hypothèse : <i>donc, d'ailleurs, également, toutefois, en effet</i>,</p>

ans. [eco-art-141-body]	<i>même, cependant, ainsi, pourtant, aussi.</i>
La modalité déontique (capital, nécessaire, important, souhaitable, utile, inutile, indispensable)	
Exemples	Commentaire
<p>À ce niveau, il est souhaitable que les experts interviennent déjà dans la construction de l'arborescence (Saaty, 1984) [geo-art-111-body]</p> <p>De ce point de vue, il est important que les recherches empiriques futures progressent dans l'analyse jointe des dynamiques des employeurs et des carrières individuelles. [eco-art-22-conclusion]</p> <p>Ensuite, le phénomène de l'ambiguïté lexicale est caractéristique des langues et, si l'on s'inscrit dans une perspective où les applications en TAL doivent demander une adaptation minimale de la part des utilisateurs, alors il apparaît nécessaire que ces applications soient capables de traiter l'ambiguïté lexicale, ce qui suppose une modélisation préalable. [tal-art-626-introduction]</p>	<p>La construction donne l'ordre, exprime une nécessité. Sa fréquence relative est plus élevée dans l'introduction où elle justifie la légitimité de la recherche.</p> <p>Le complément phrastique est en subjonctif.</p> <p>Le verbe copule est le plus souvent à l'indicatif présent. Néanmoins, il apparaît également au conditionnel présent et à l'imparfait.</p>
La sous-costruction : Il + V + ADJ + que + subj.	
L'évaluation (étonnant, rare, normal, absurde, acceptable, banal, bon, commode, commun, difficile, erroné, excessif, fréquent, fructueux, inintéressant intéressant, légitime, mauvais, optimal, piquant, radical, significatif, théorique)	
Exemples	Commentaires
Il n'est guère étonnant que les	La construction permet de juger les

<p>personnages masculins du Satiricon ne soient pas en toge : le texte de Pétrone ne montre pas un monde de citoyens dans l'exercice de leurs fonctions. [his-art-127-notes]</p> <p>Il est significatif que la thèse de 1892 se termine par une collection d'auto-observations souvent assez longues classées par ordre alphabétique d'auteur et qui constituent comme une sorte de caution prestigieuse, voire d'argument de vente. [lin-art-23-body]</p> <p>Il est significatif que pour photographier la société française du début des années quatre-vingt, un organisme aussi officiel que la Datar (Délégation à l'aménagement du territoire et à l'activité régionale) fasse appel à des photographes d'art et non à des documentaristes [Paysages, photographies. La Mission photographique de la Datar, 1985]. [ant-art-265-body]</p>	<p>phénomènes, les événements en présentant son jugement sous une forme de construction impersonnelle. Elle permet aussi de souligner l'importance d'un événement, d'attirer l'attention sur un événement. En ce qui concerne la fonction discursive, elle permet d'introduire un nouveau thème.</p> <p>Le verbe copule est à l'indicatif présent, au conditionnel présent, au futur simple, à l'imparfait.</p> <p>Le complément phrastique est au subjonctif.</p>
--	--

2. *Il + V + ADJ + de*

La structure	Il + V + ADJ + de + inf.
La définition	<p>Il [sujet apparent] + verbe copule + adjectif [attribut du sujet] + de + infinitif</p> <p>Une construction impersonnelle à sujet phrastique extraposé, composée du pronom impersonnel <i>il</i> suivi d'un verbe copule, un adjectif et la préposition <i>de</i> qui introduit le verbe à l'infinitif.</p>

Les constituants et leurs propriétés morphologiques	<p>Il – pronom impersonnel, m, sg</p> <p>Verbe, 3 per, sg, indicatif</p> <p>Adjectif, m, sg</p> <p>de – préposition</p>
La tête lexicale	<p>ADJ : économique, égal, éloquent, équivalent, étonnant, étrange, évident, absent, absurde, agréable, aisé, aléatoire, analysable, anodin, arbitraire, avantageux, banal, beau, bon, capital, caricatural, clair, cohérent, commode, commun, congruent, constant, consubstantiel, correct, crucial, curieux, délicat, différent, difficile, douteux, effectif, envisageable, erroné, essentiel, exact, facile, faisable, fallacieux, faux, fondamental, fréquent, hâtif, habituel, hasardeux, homogène, idiot, illégal, illogique, illusoire, immédiat, impensable, important, impossible, imprudent, inévitable, inconcevable, incontestable, indéniable, indifférent, indirect, indispensable, inexact, influent, inintéressant, injuste, instructif, inutile, invraisemblable, juste, légitime, logique, loisible, long, minimal, modeste, nécessaire, négligeable, naturel, normal, notable, notoire, observable, optimal, paradoxal, pertinent, piquant, plausible, positif, possible, préférable, prématuré, présent, primordial, probable, productif, profitable, réducteur, révélateur, radical, raisonnable, rare, rationnel, remarquable, sage, significatif, souhaitable, spécifique, superflu, surprenant, susceptible, symptomatique, tolérable, typique, urgent, usuel, utile, utopique, vain, valable, vital, vrai, vraisemblable</p> <p>VERB : être, sembler, devenir, paraître</p>
LES FONCTIONS	
Modalité épistémique : probabilité (possible, impossible, envisageable, plausible)	
Exemples	Commentaire
Nous avons montré que, sous certaines conditions, il était possible de passer de la répartition à la capitalisation de façon à	La construction indique la probabilité de réalisation. Elle sert principalement à introduire l'hypothèse et exprimer les

<p>accroître l'utilité de tous. [eco-art-385-conclusion]</p> <p>Il est donc possible de négliger toutes ces variations, non seulement en raison de leur faible ampleur, mais encore parce qu'elles se neutralisent ; aussi sera-t-il possible de travailler, dans l'approche du phénomène du jugement de paraphrase dans le corpus, sur des nombres plus grands que si l'on prenait toutes les catégories séparément. [sed-the-317-body]</p> <p>De fait, il nous est impossible de tester la médiation de l'effet du conflit et des buts sur l'apprentissage, par le mode de régulation du conflit. [psy-the-504-body]</p>	<p>restrictions et les possibilités de la recherche.</p> <p>La construction est très souvent accompagnée par un adverbe ou une locution adverbiale qui diminue ou renforce la probabilité tel que <i>pratiquement, sans doute, peut-être, peu, totalement, tout à fait, certes</i>, l'adverbe qui restreint la possibilité en ce qui concerne sont aspect temporel : <i>actuellement, souvent, parfois, rarement, toujours, pas encore</i> et/ou un adverbe, une conjonction qui introduit ou justifie l'introduction de l'hypothèse : <i>en effet, également, donc, en revanche, alors, tout de même, cependant, par conséquent, néanmoins, d'ailleurs, encore, ainsi</i>.</p> <p>Elle contient très souvent le pronom personnel <i>nous</i> ou, plus rarement, <i>leur, lui</i></p> <p>Les verbes copules sont : <i>être, devenir, sembler, rester, paraître, apparaître</i>. Ils sont soit au présent de l'indicatif, soit au conditionnel présent.</p>
<p>La modalité déontique (nécessaire, important, indispensable, essentiel, souhaitable, crucial, vital)</p>	
Exemples	Commentaire
<p>C'est pourquoi il est crucial de définir les conditions de possibilité d'une enquête fondée sur le rapport spéculaire entre un groupe professionnel et sa visibilité médiatique. [incom-art-5-body]</p>	<p>La construction exprime une nécessité. Elle permet très souvent au chercheur d'introduire un thème d'analyse, d'expliquer son procès de raisonnement, de justifier ses choix lors d'analyse. C'est pourquoi la construction peut être trouvée</p>

<p>Il est important de noter que les y sont orientés vers le bas (une trace plus haute est caractérisée par des y plus petits). [psy-the-419-body]</p> <p>Les éléments de comparaison entre les savoirs théoriques et les savoirs d'action montrent que ces savoirs se distinguent sur le plan théorique. Il est donc nécessaire de les différencier pour des raisons analytiques. [sed-the-356-body]</p>	<p>dans tout l'article sans préférence pour une partie particulière.</p> <p>La construction est très souvent accompagnée par un adverbe ou une locution adverbiale qui justifie, explique la nécessité : <i>aussi, en effet, donc, toutefois, également, alors, en outre, ensuite, pourtant, par conséquent</i> ou souligne son importance : <i>particulièrement, très, tout à fait</i></p> <p>L'infinitif qui suit la préposition <i>de</i> concerne très souvent le modèle d'analyse, il démontre alors un thème d'analyse : <i>se centrer sur, prendre en compte, interroger</i>, ou il concerne l'analyse elle-même, à savoir la description des données (en se focalisant très souvent sur un aspect particulier) ou l'interprétation des résultats : <i>approfondir, compléter, expliciter, effectuer, observer, analyser, expliquer, rappeler, etc.</i></p> <p>Le verbe copule utilisé dans cette construction est uniquement <i>être</i>. Le plus souvent, il est à l'indicatif présent ; pourtant, nous le pouvons aussi rencontrer au conditionnel présent, au futur simple ou à l'imparfait.</p>
<p>La valorisation (difficile, facile, utile, aisé, préférable, rare, pertinent, délicat, vain, légitime, inutile, optimal, illusoire, loisible, économique, naturel, erroné, étonnant, révélateur, absurde, profitable, logique, fréquent, hâtif, hasardeux, juste, bon, faux, absent, normal, réducteur, idiot, symptomatique, instructif, commode, immédiat, injuste, paradoxal, usuel, indifférent, correct, cohérent, illégal, anodin, utopique, prématuré, banal,</p>	

inintéressant, radical, piquant, remarquable, impensable, aléatoire, inexact, illogique, exact)	
Exemples	Commentaires
<p>Il est difficile de dire combien de textes exactement sont concernés ; il faudrait pour cela fixer une limite, arbitraire, au-delà de laquelle décider que l'usage des marques personnelles et du <i>il</i> impersonnel ou du passif est à considérer comme particulièrement élevé. [lin-the-115-body]</p> <p>En premier lieu, il est utile de faire le point sur les défauts structurels de la gestion publique française. [his-art-166-body]</p> <p>Il est prématuré de prétendre, avec les seuls éléments analysés, proposer une hypothèse.</p>	<p>La construction permet de juger les phénomènes, les évènements en présentant son jugement sous une forme de construction impersonnelle. Outre cela, elle remplit une fonction discursive en permettant d'introduire un nouveau thème.</p> <p>Nous y voyons les adjectifs axiologiques (<i>pertinent, logique, utile</i>), portant sur la temporalité (<i>rare, fréquent</i>), sur la complexité (<i>difficile, facile</i>). Pourtant, la plupart d'entre eux sont des adjectifs évaluatifs qui n'entrent pas dans des catégories plus spécifiques : <i>correct, utopique, banal, loisible</i>.</p> <p>La construction est parfois accompagnée par le pronom personnel <i>nous</i>.</p> <p>Le verbe copule (<i>être, sembler, devenir, rester, paraître, apparaître</i>) est le plus souvent au présent de l'indicatif, au conditionnel présent ou à l'imparfait.</p>

3. *Il + V + que + phrase*

La structure	Il + V + que + phrase
La définition	<p>Il [sujet apparent] + verbe + que [conjonction] + phrase</p> <p>C'est une construction impersonnelle qui se compose du pronom impersonnel <i>il</i> (le sujet apparent) suivie d'un verbe à la troisième</p>

	personne du singulier et la conjonction de subordination <i>que</i> qui introduit la subordonnée complétive.
Les constituants et leurs propriétés morphologiques	Il – pronom impersonnel, m, sg Verbe, 3 per, sg, indicatif Que - conjonction
La tête lexicale	échapper, apparaître, arriver, s'avérer, convenir, découler, s'ensuivre, importer, manquer, paraître, se pouvoir, résulter, ressortir, rester, sembler, sortir, suffire
LES FONCTIONS	
Etat : existence (sembler, rester, pouvoir, paraître, apparaître, s'avérer, arriver)	
Exemples	Commentaire
<p>Il reste que la revendication de Gasc pour une planification raisonnée se situe à une jointure de la rationalité éducative qui n'a plus jamais été abolie, quand bien même elle est la cause de nos tourments. [sed-art-3-body]</p> <p>Plus généralement, il semblerait que le critère d'identité de fonctionnement structural soit efficace pour circonscrire celles des variables phonologiques que nous avons appelées symétriques. [lin-hdr-396-body]</p> <p>Plus subtilement, il se peut que le fait même de débattre une question puisse ébranler les certitudes initiales de protagonistes. [psy-hdr-647-body]</p>	<p>La construction sert à informer de l'existence d'un fait, avec plus (<i>rester, arriver, s'avérer</i>) ou moins (<i>sembler, se pouvoir, apparaître, paraître</i>) de certitude. La construction remplit aussi une fonction discursive : elle permet d'introduire un nouveau thème.</p> <p>La construction est fréquente dans le développement et dans les conclusions de l'article.</p> <p>Le verbe est le plus souvent à l'indicatif présent, au conditionnel présent ou à l'imparfait.</p> <p>Le complément phrastique est soit à l'indicatif (<i>rester, demeurer</i>), soit au subjonctif (<i>arriver, sembler, se pouvoir, paraître, falloir</i>).</p>

Etat : conclusion (s'ensuivre, ressortir, résulter, découler)	
Exemples	Commentaire
<p>Il s'ensuit que les séries télévisées constituent des vecteurs majeurs pour connaître la police ; il s'agit d'un savoir collectif acquis par procuration, c'est-à-dire par médias interposés. [infcom-art-5-body]</p> <p>Il en découle que le programme d'optimisation de l'entreprise peut s'écrire. [eci-art-221-body]</p> <p>Il ressort de ces travaux que les comportements d'adhésion apparaissent comme fortement contraints par le contexte interactionnel dans la mesure où ces comportements ne sont pas introduits par l'enfant à n'importe quel moment de l'interaction mais consécutivement à certaines sollicitations du parent [psy-art-382-body]</p>	<p>La construction permet de tirer des conclusions de ce qui a été dit, d'un fait présenté. C'est pourquoi la construction est fréquente dans des conclusions.</p> <p>La cause de la conclusion peut être exprimée au niveau de la construction même grâce au syntagme nominal introduit à l'aide de la préposition <i>de</i>. Dans le cas contraire, il arrive souvent que la construction contienne le pronom anaphorique <i>en</i>.</p> <p>Le complément phrastique est à l'indicatif.</p>
La modalité déontique (importer, falloir, convenir)	
Exemples	Commentaires
<p>Il importe seulement que les données pragmatiques nécessaires à l'interprétation des énoncés ne court-circuitent pas les éléments que peut fournir l'analyse linguistique esquissée ci-dessus à propos de d'ailleurs. [lin-art-176-body]</p> <p>Cependant, il faut également que les traits distinctifs de la cible soient suffisamment saillants pour pouvoir être extraits de la</p>	<p>La construction permet d'exprimer les nécessités méthodologiques de la recherche ou les conditions nécessaires pour qu'une hypothèse soit vraie. La construction est absente dans l'introduction.</p> <p>La construction admet un adverbe ou une locution adverbiale : <i>néanmoins, aussi, à tout le moins, donc, seulement, d'abord,</i></p>

<p>stimulation. [psy-art-442-body]</p> <p>Conformément à notre hypothèse, il faut que leur identité sociale soit contextuellement menacée et dévalorisée pour que les individus cherchent à protéger leur estime de soi, via l'identification à leur groupe menacé. Autrement dit, il faut que la performance des autres membres du groupe soit perçue comme étant susceptible de remettre en cause le bien-fondé de l'image négative de leur groupe.</p>	<p><i>également.</i></p> <p>La construction peut contenir aussi le résultat souhaité de l'injonction introduit par la préposition <i>pour</i>.</p> <p>Le complément phrastique est au subjonctif.</p>
---	---

4. *Il + V + de + inf.*

La structure	Il + V + de + inf.
La définition	Il [sujet apparent] + verbe + de [préposition] +inf. C’est une construction impersonnelle qui se compose du pronom impersonnel <i>il</i> (le sujet apparent) suivie d’un verbe à la troisième personne du singulier et la préposition <i>de</i> qui introduit une proposition infinitive.
Les constituants et leurs propriétés morphologiques	Il – pronom impersonnel, m, sg Verbe, 3 per, sg, indicatif de - préposition
La tête lexicale	Importer, suffir, convenir, s’agir, sembler
LES FONCTIONS	
Modalité déontique (importer, suffir, convenir, s’agir)	
Exemples	Commentaire
Mais avant d’étudier les dynamiques actuelles, il convient de retracer les	La construction sert à exprimer les nécessités méthodologiques ou des

<p>différentes étapes de construction urbaine afin de comprendre les spécificités des processus en cours. [geo-art-7-body]</p> <p>Si on accorde aux discours à visée scientifique (Granger, 1993) sur le social le même statut qu'aux discours sur la nature, le travail transpositif s'impose pour les enseignements économiques et sociaux, au même titre que pour les disciplines scientifiques. Il importe alors d'en étudier les spécificités. [sed-art-49-body]</p> <p>L'intérêt de la sociologie interactionniste est de ne pas rabattre mécaniquement la réception sur une appartenance sociale donnée une fois pour toutes : puisque la profession n'est pas écrasement mais négociation du sens, il s'agit d'observer, dans le rapport spéculaire aux médias, tant les déterminations exercées par l'appartenance professionnelle que le sens élaboré et partagé sur cette même appartenance (Morley, 2006). [infcom-art-5-body]</p>	<p>approfondissements de la recherche suggérés pour des recherches ultérieures. La conclusion est la partie de l'article ou la construction est relativement plus fréquente.</p> <p>Le verbe est le plus souvent à l'indicatif présent, au conditionnel présent, au futur simple ou à l'imparfait.</p> <p>L'infinitif qui suit la préposition <i>de</i> est un verbe d'analyse (<i>montrer, établir, examiner, analyser, déterminer, construire, mesurer</i>).</p> <p>La construction peut contenir des adverbes qui soulignent la nécessité ou qui introduisent une suite logique de ce qui a été dit comme <i>alors, certainement, d'abord, avant tout, donc, enfin, premièrement, simplement</i>.</p>
--	--

5. Pouvoir + être + pp

La structure	pouvoir + être + pp
La définition	<p>Pouvoir [verbe modal] + être + participe passé</p> <p>C'est une construction passive qui se compose du verbe <i>pouvoir</i>, suivi du verbe être, l'auxiliaire d'un participe passé.</p>

Les constituants et leurs propriétés morphologiques	Pouvoir – verbe modal conjugué Etre – l’infinitif Participe passé
La tête lexicale	Verbe – pouvoir Participe passé - associer, assimiler, lier, ramener, réduire, généraliser, comparer, rapporter, relier, corréler, conduire, considérer, interpréter, voir, décrire, définir, analyser, lire, percevoir, comprendre, envisager, appréhender, concevoir, traiter, penser, représenter, saisir, entendre, désigner, utiliser, prendre en compte, faire, réduire, compléter, réaliser, étudier, adapter
LES FONCTIONS	
Fonction : comparaison, association (associer, assimiler, lier, ramener, réduire, généraliser, comparer, rapporter, relier, corréler, conduire)	
Exemples	Commentaire
<p>Une partie des écarts peut être liée à la plus grande proportion de femmes (de longévité supérieure) dans la catégorie des bas salaires. [eco-art-9-body]</p> <p>Si les phénomènes collectifs peuvent être ramenés à leurs composantes individuelles et que celles-ci sont déterminées par des facteurs infra-individuels, il faudra bien, selon une logique de transitivity, éclairer celles-là par celles-ci [soc-art-60-body]</p> <p>Mais ce résultat pourrait être simplement dû à un manque de précision dans les estimations, par exemple, à cause d’erreurs de mesure sur l’offre de travail. [eco-art-132-notes]</p>	<p>La construction permet de mettre en rapport deux phénomènes, le premier étant très souvent la conséquence du deuxième.</p> <p>Le verbe <i>pouvoir</i> est le plus souvent à l’indicatif présent. Il apparaît aussi au conditionnel présent, au futur simple et à l’imparfait.</p> <p>Les participes passé proviennent des verbes d’association ou des verbes de cause.</p>

LES FONCTIONS	
Fonction : interprétation (considérer, interpréter, voir, décrire, définir, analyser, lire, percevoir, comprendre, envisager, appréhender, concevoir, traiter, penser, représenter, saisir, entendre, désigner)	
Exemples	Commentaire
<p>Réciproquement, tout procès de sémiotisation pourra être abordé comme un <i>cours d'action</i>, rassemblant plusieurs protagonistes (éventuellement " fictifs "), captés par un certain " jeu " ou " genre " sémiotique. [ant-art-121-body]</p> <p>La paramétrisation de l'utilité marginale de la consommation peut être considérée comme une manière de capturer une certaine forme d'hétérogénéité observable du facteur d'escompte des ménages ou de leur degré de patience. [eco-art-132-body]</p> <p>La place croissante du paysage dans l'aménagement forestier pourrait être lue comme le signe d'un manque de fonctionnalités de ces territoires boisés à l'échelle nationale. [geo-art-393-body]</p>	<p>La construction permet d'estimer, de juger un phénomène en le comparant à un autre. Elle sert très souvent à définir, à préciser un terme, à tirer des conclusions. Présente dans toutes les parties de l'article, elle apparaît le plus fréquemment dans la conclusion et dans le développement.</p> <p>Le verbe <i>pouvoir</i> et <i>devoir</i> est le plus souvent à l'indicatif présent. Ils apparaissent aussi au conditionnel présent, au futur simple et à l'imparfait.</p> <p>Les participes passés proviennent des verbes de compréhension, d'interprétation ou d'évaluation.</p>

6. devoir + être + pp

La structure	devoir + être + pp
La définition	<p>Devoir [verbe modal] + être + participe passé</p> <p>C'est une construction passive qui se compose du verbe <i>devoir</i>, suivi du verbe être, l'auxiliaire d'un participe passé.</p>

Les constituants et leurs propriétés morphologiques	devoir – verbe modal conjugué Etre – l’infinitif Participe passé
La tête lexicale	Verbe – devoir Participe passé - utiliser, prendre en compte, faire, réduire, compléter, réaliser, étudier, adapter
LES FONCTIONS	
Fonction : examen, démonstration (utiliser, prendre en compte, faire, réduire, compléter, réaliser, étudier, adapter)	
Exemples	Commentaire
<p>Un effort intellectuel important et un positionnement épistémologique rigoureux doivent être menés par chacun: suis-je praticien-enseignant, formateur ou chercheur lorsque je tiens tel ou tel propos? [sed-art.-104-body]</p> <p>Bien entendu, le lien entre les deux phénomènes doit être étudié de manière plus détaillée et vérifié à partir des représentations exprimées par les contemporains. [his-art.-135-body]</p> <p>Il est cependant clair que l’étude présentée ici doit être complétée par les cas où la satellite d’une relation subordonnante précède le noyau. [tal-art.-236-body]</p>	<p>La construction permet de démontrer les nécessités de l’étude. Elle est relativement plus fréquente dans l’introduction où elle démontre les enjeux de la recherche présentés dans l’article.</p> <p>Le verbe modal est à l’indicatif présent, au conditionnel présent, au futur simple, à l’imparfait.</p> <p>Les participes passés proviennent des verbes d’examen, de réalisation et de démonstration.</p>

Table des matières

Remerciements	4
Sommaire	7
Introduction	8
PARTIE 1 - L'ECRIT SCIENTIFIQUE	11
CHAPITRE 1. GENRE ET ECRITS SCIENTIFIQUES	12
1. Variation et genre	12
2. Langue generale et langue de specialite	14
3. L'article scientifique comme un genre	16
CHAPITRE 2. CARACTERISTIQUES LINGUISTIQUES DE L'ECRIT SCIENTIFIQUE	19
1. Genre et l'analyse syntaxique	19
2. Caractéristiques du discours scientifique selon des travaux précédents	21
2.1. Aspects lexicologiques	22
2.2. Aspects syntaxiques et grammaticaux	23
2.3. Aspects textuels - cohérence textuelle	26
3. Constructions syntaxiques	27
3.1. Le lexique-grammaire	28
3.2. Les patrons verbaux de l'écrit scientifique - recherche de Yan (2013)	29
3.3. Grammaires de construction	31
4. Le modèle de construction adopté	34
PARTIE 2 - METHODOLOGIE	37
CHAPITRE 3. LA RECHERCHE DES CONSTRUCTIONS	38
1. Scientext	38
2. Le choix des corpus	39
3. Le traitement du corpus	39
4. Les résultats	41
1. Le corpus de contraste	42
PARTIE 3 - LES RESULTATS ET L'OBSERVATION DES CONSTRUCTIONS	45
CHAPITRE 4. LES RESULTATS QUANTITATIFS	46
CHAPITRE 5. LA DESCRIPTION DES CONSTRUCTIONS	51
1. L'impersonnel	51
1.1. Il + V + ADJ + que	51
1.2. Il + V + ADJ + de	53
1.3. Il + V + que	55
1.4. Il + V + de	57
2. Le passif	58
2.1. Pouvoir + être + pp	59
3. Conclusion	60
Conclusion	63
Bibliographie	65
Glossaire	67
Sigles et abréviations utilisés	68
Table des illustrations	69
Table des annexes	70
Table des matières	94

MOTS-CLÉS : l'écrit scientifique, l'analyse du genre, grammaires de construction, linguistique du corpus

RÉSUMÉ

Les langues de spécialité sont traditionnellement analysées du point de vue de leur terminologie. À l'inverse de cette tendance générale, le présent travail se veut concentrer sur l'aspect grammatical de l'écrit scientifique en ayant recours aux théories de lexique-grammaire et des grammaires de construction. Après avoir repéré les constructions grâce au corpus Scientext et la comparaison de la fréquence relative de ces constructions dans le corpus de contraste, on a essayé de décrire quelques-unes d'elles à l'aide du modèle élaboré.

KEYWORDS : scientific writing, genre analysis, grammar constructions, corpus linguistics

ABSTRACT

Speciality languages are traditionally analyzed from the point of view of their terminology. Contrary to this general tendency, the present work aims to focus on the grammatical aspect of scientific writing through the use of lexicon-grammar theory and the use of construction grammars. After discovering the constructions through Scientexte corpus and the comparison of the relative frequency of those constructions in the contrast corpus, we tried to describe some of them using the developed model.

Titre de l'illustration de couverture, source, date, voire commentaire

